The Dissertation Committee for Sanjeev Vinayak Namjoshi Certifies that this is the approved version of the following dissertation:

Bioinformatic Approaches to Screening the Molecular Framework Underlying Local Dendritic mRNA Translation

Committee:

Johann Hofmann Supervisor

Kimberly F. Raab-Graham Co- Supervisor

Nigel S. Atkinson

Edward M. Marcotte

Christopher S. Sullivan

Bioinformatic Approaches to Screening the Molecular Framework Underlying Local Dendritic mRNA Translation

by

Sanjeev Vinayak Namjoshi, B.S.Bio

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2017

Dedication

I dedicate this work to all those who have been patient and supportive of me for over half a decade. To Dr. Kimberly Raab-Graham for her guidance and vision which has nurtured me into the scientist I am today. To the Raab-Graham lab for listening to my rambling discourse, bad music, and being a clarifying force in moments of confusion. To my mother for her unending supply of wisdom and council so I could better understand myself. To my father for questioning my own assumptions so that I could become a better scientist. To my sister for helping me realize what it is I believe. To my friends: Ariana, Alex & Alex, Jennifer, Kalen, Leslie, Peter, and Rick, who always inquired about my progress and emboldened me to move forward. To Elissa Jo for her constant encouragement and unwavering optimism at the most difficult moments. Finally, I dedicate the following quote to all the rodents whose heads have been sacrificed for my *projects*:

"The world to which we have belonged offers nothing to love outside of each individual insufficiency: its existence is limited to utility […] Existence is not only an agitated void, it is a dance that forces one to dance with fanaticism. Thought that does not have a dead fragment as its object has the inner existence of flames […] Human life is exhausted from serving as the head of, or the reason for, the universe. To the extent that it becomes this head and this reason, to the extent that it becomes necessary to the universe, it accepts servitude. If it is not free, existence becomes empty or neutral and, if it is free, it is in play. The Earth, as long as it only gave rise to cataclysms, trees, and birds, was a free universe; the fascination of freedom was tarnished when the Earth produced a being who demanded necessity as a law above the universe. Man, however, has remained free not to respond to any necessity; he is free to resemble everything that is not himself in the universe. He can set aside the thought that it is he or God who keeps the rest of things from being absurd. Man has escaped from his head just as the condemned man has escaped from his prison […] Beyond what I am, I meet a being who makes me laugh because he is headless […] The fate and the infinite tumult of human life is open not to those who exist like gouged out eyes, but to those who are like seers, carried away by a shattering dream that could not belong to them."

-Georges Bataille, *La Conjuration Sacree* (1937)

Acknowledgements

I would like to acknowledge members of the Raab-Graham lab, past and present, who have helped me along the way and contributed to my continued success and dedication: Chelcie Heaney, Emily Workman, Farr Niere, Luisa Cacheaux, Natasha Sosanya, Sarah Wolfe, and all rotation students and undergrads that have passed through the lab the last 3 and a half years. Thanks to Dr. Hans Hofmann his mentorship and advice. Thanks to my committee for their guidance and helpful discussion. And finally, thanks to Dr. Kimberly Raab-Graham for taking a chance on me. Thank you for all you have done.

Bioinformatics Approaches to Screening the Molecular Framework Underlying Local Dendritic mRNA Translation

Sanjeev Vinayak Namjoshi, Ph.D

The University of Texas at Austin, 2017

Supervisor: Johann Hofmann

Co-Supervisor: Kimberly F. Raab-Graham

Learning and memory systems within the brain are believed to be the result of molecular events that occur at the synapse. These molecular events control synaptic efficacy – the modulation of the underlying molecular architecture of the synapse to influence the strength or weakness of connections between pre- and postsynaptic neurons. Deficits in any of these molecular processes results in neuronal dysfunction. The Mammalian/Mechanistic Target of Rapamycin Complex (mTORC1) is a protein complex within the brain composed of the serine/threonine kinase mTOR and other interacting proteins. Dysregulation of mTOR disrupts the many processes under its control such as local dendritic translation at the synapse. Since mTOR is at the core of many important signaling pathways, aberrant mTOR activity results in neuronal disease. The network of interactions between these molecular components is vast, forming an interconnected system that is dynamic and directed. In order to better understand the mechanistic nature of these interactions, the application of high-throughput technologies must be employed. Here we utilize multiple bioinformatics approaches combined with high-throughput technologies to clarify the role of mTOR in local dendritic translation. Using mass spectrometry, we provide the first evidence that mTOR bidirectionally controls the

expression of over 700 proteins in the cortex, many of which are known to be associated with diseases in which mTOR is overactive (Chapter 2). Our investigations reveal a novel role PARK7 in tuberous sclerosis complex. Furthermore, we use and develop multiple bioinformatics tools to further delineate the nature of the RNA-binding properties of PARK7 and alternative avenues for drug discovery (Chapter 3). Finally, we provide evidence that the Fragile-X Mental Retardation protein sequesters a population of mRNAs involved in trans-synaptic signaling that mTOR translates to remodel the synapse, providing a mechanistic basis for the action of rapidly-acting antidepressants (Chapter 4). Collectively, our work stresses the importance of applying high-throughput technologies to answer long-standing questions in the field of local dendritic translation. Our findings provide new avenues of investigation and research to better understand neuronal disease and synaptic plasticity.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction[1]

LOCAL TRANSLATION AND DENDRITIC RNA TRANSPORT

Long-term memory formation relies on the modulation of synaptic efficacy - the strengthening or weakening of connections between a presynaptic and postsynaptic cell. Such changes are dependent on the alteration of the underlying neuronal architecture of the synapse through protein synthesis in the dendrites. Many of these changes are controlled by locally synthesized protein kinases that alter ion channel densities and conductivities to change the strength of connections between neurons (Martin et al., 2000; Giese, 2012; Papoutsi et al., 2012; Stewart and Popov, 2012). In order for the changes made at the synapse to be long-lasting and consolidated, proteins must be synthesized rapidly at the base of a region of the neuron called the dendritic spine Thus, constitutive and activity regulated mRNA trafficking in neuronal cells allows localized protein synthesis in specific compartments or areas of the neuron far from the soma such as axons, dendrites, and spines (Jung and Holt, 2011). Consequently, many RNA transcripts coding for proteins that induce changes in synaptic efficacy are localized in dendrites and ready for rapid expression through local mRNA translation.

Historically, all mRNAs were thought to be exclusively translated in the soma. This dogmatic view was questioned when observations in a series of electron microscopy (EM) studies revealed the presence of polyribosomes in dendrites, specifically at the base of dendritic spines of the dentate gryus (Bodian, 1965; 1972; Peters et al., 1976). It was not until 1983 that Steward and Levy provided the first quantitative evidence of synapse-associated polyribosome complexes and their localization (Steward and Levy, 1982). Steward and Levy hypothesized that synapse-associated polyribosome complexes may be

[1] Parts of this chapter are published in Namjoshi, S. and Raab-Graham, K.F. (2017). Screening the Molecular Framework Underlying Local Dendritic mRNA Translation. Front. Mol. Neurosci. 10(45). doi: 10.3389/fnmol.2017.00045. I wrote the manuscript and created all figures and tables.

necessary for the expression of proteins that constitute the synapse due to their proximity to dendritic spines. Their hypothesis was confirmed through numerous studies in the two decades that followed demonstrating the requirement for local protein synthesis in processes related to synaptic plasticity and learning. Some of these early studies demonstrated correlations between polyribosome numbers and synaptogenesis suggesting that the synapse-associated polyribosome complexes were the source of the proteins found in the postsynaptic density (PSD) (Steward and Falk, 1985; 1986; Palacios-Pru et al., 1988). This was followed by a number of key studies that identified select mRNA transcripts that were localized and translated in the dendritic spines (Steward et al., 1996; Steward and Schuman, 2001). Among these was the important discovery that BDNF-induced synaptic potentiation required local protein synthesis (Kang and Schuman, 1996).

Dendritic mRNA transport relies on complex formation of RNA granules. RNA granules contain RNA-binding proteins (RBPs) - which bind to sequestered mRNAs to inhibit their translation - as well as some translation factors, ribosomes, and other proteins that control translation (Kiebler and Bassell, 2006). Upon synaptic activation, select repressed mRNAs localized to the synapse are translated where the ribosome within the RNA granule can initiate rapid translation into the required protein product (Kim et al., 2013b; Pimentel and Boccaccio, 2014). Notably, RBPs play a vital role in learning in memory. The absence of an RBP resulting from incorrect localization or dysfunction due to mutations may lead to aberrant translation or repression of specific mRNAs under its control resulting in a neurological disorder (Sephton and Yu, 2015). Furthermore, the activity of protein kinases, such as mammalian target of rapamycin (mTOR), is coupled to translation to facilitate processes related to learning and memory. Disruption of these processes can lead to neuronal dysfunction (Giese and Mizuno, 2013; Lipton and Sahin, 2014). Many animal disease models that reproduce both symptoms and genetic alterations

seen in humans show dysregulated local mRNA translation (Pei and Hugon, 2008; Zang et al., 2009; Sharma et al., 2010; Ricciardi et al., 2011; Devi and Ohno, 2013; Ma et al., 2013).

mTOR AND THE CONTROL OF LOCAL TRANSLATION

Long-lasting changes to synaptic plasticity require N-methyl-D-aspartate glutamate receptor (NMDAR) activation (Graber et al., 2013b). The resulting signaling cascade activates mammalian/mechanistic target of rapamycin (mTOR) via the PI3K-TSC pathway (outlined below, see Figure 1.1). The local translation process is largely mediated by the mammalian/mechanistic target of rapamycin (mTOR). mTOR is a 289-kDa serine/threonine protein kinase that plays a central role in the regulation and balance of a large number of biological processes including the regulation of dendritic protein synthesis. Essentially a master regulator cell growth, mTOR responds to a large variety of upstream signals and processes including stress (low energy, oxygen depletion, DNA damage), nutrient sensing/availability, and response to growth factors (Betz and Hall, 2013). Many processes are modulated downstream of mTOR signaling, most notably the control of protein synthesis/homeostasis, but also lipid synthesis, autophagy, hormone regulation, and synaptic plasticity involved in learning/memory circuitry. Due to its central role in many cellular processes it is not surprising that mTOR is also involved in crosstalk with numerous other pathways including PI3K/Ras (via TSC1/2), growth factors (IGF1, Akt, Erk), energy (AMPK), oxygen sensing (REDD1), and DNA damage pathways (Laplante and Sabatini, 2009; Wang and Proud, 2011). Due to the vast number of signals mTOR responds to and the processes it modulates, a greater understanding of the protein and functional landscape altered by mTOR activity is vital for determining the causes of a number of diseases associated with mTOR dysfunction.

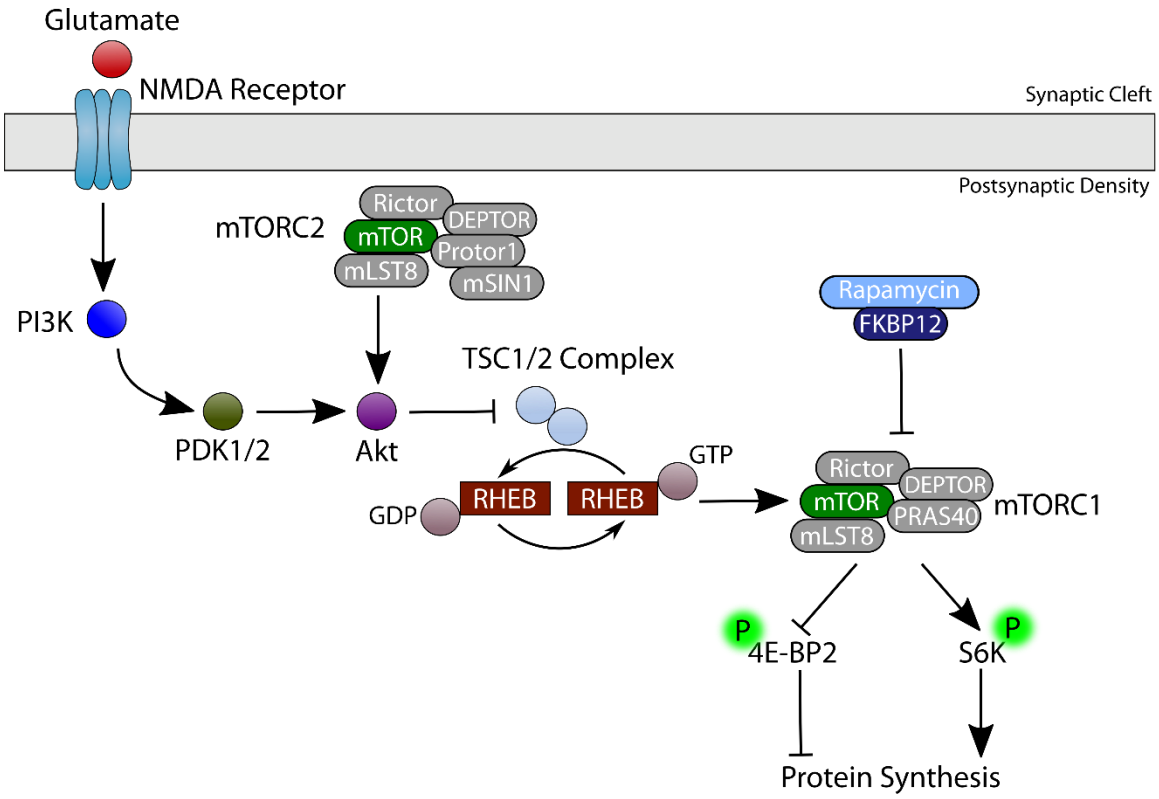


Figure 1.1: NMDA receptor-mediated activation of mTORC1

Glutamate or glycine binding to NMDA receptors on the post-synaptic side of the neuron triggers a signaling cascade that activates mTORC1. Signaling through PI3K and PDK1/2 activates Akt. Akt blocks the tuberous sclerosis complex (TSC1/2) complex, an important intersection point for pathway crosstalk. mTORC1 is typically activated through the binding of small GTP-binding protein Ras homolog enriched in brain (RHEB). RHEB's GDP/GTP states are controlled by TSC1/2, a RHEB GTPase-activating protein and heterodimer. When it activates RHEB, in its GTP-bound state, RHEB is able to activate mTOR. mTORC1 has five known components. These include the catalytic mTOR subunit, regulatory-associated protein of mammalian target of rapamycin (RAPTOR), mammalian lethal with Sec13 protein 8 (mLST8), proline-rich Akt substrate 40 kDa (PRAS40), and DEP-domain containing mTOR-interacting protein (DEPTOR). mTORC1 then regulates S6K and 4E-BP2 which regulate translation in specific contexts. mTORC2 enters this pathway through the activation of Akt. Although mTORC1 and mTORC2 share a number of proteins in common, they respond to a different set of cellular signals. Once inside the cell, rapamycin binds to FK506-binding protein of 12 kDa (FKBP12). This complex then interacts with the rapamycin binding domain (FRB) of mTOR and inhibits mTORC1 function and activities in various processes.

mTOR's ability to respond to various signals and carry out its functions is dependent on the formation of two different protein complexes referred to as mTORC1 (mTOR complex 1) and mTORC2 (mTOR complex 2). We will focus on mTORC1 for the purpose of this introduction. mTORC1 has five known components. These include the catalytic mTOR subunit, regulatory-associated protein of mammalian target of rapamycin (RAPTOR), mammalian lethal with Sec13 protein 8 (mLST8), proline-rich Akt substrate 40 kDa (PRAS40), and DEP-domain containing mTOR-interacting protein (DEPTOR) (Vander Haar et al., 2007; Peterson et al., 2009). The specific roles of the proteins in each complex has not been well-characterized. In the case of mTORC1, Raptor may regulate complex assembly by recruiting mTOR substrates while PRAS40/DEPTOR has been observed to negatively regulate mTORC1 when its activity levels decline (Hara et al., 2002; Kim et al., 2002). Although mTORC1 and mTORC2 share a number of proteins in common, they respond to a different set of cellular signals and activate different downstream targets.

mTORC1 activation and rapamycin inhibition

Typically, "activation" of mTORC1 is measured by its downstream output. mTORC1 is known to play an important role in the regulation of translation and so mTORC1 activation is usually measured through levels of S6 kinase 1 (S6K1), its substrates, or 4EBP1 – a translation initiation factor. mTORC1 is typically activated through the binding of small GTP-binding protein Ras homolog enriched in brain (RHEB). RHEB's GDP/GTP states are controlled by the tuber sclerosis complex (TSC), a RHEB GTPase-activating protein and heterodimer consisting of TSC1 and TSC2. When it activates RHEB, in its GTP-bound state, RHEB is able to activate mTOR (Huang and

Manning, 2008). The direct mechanism of mTOR activation by RHEB is unclear however, it may be through direct binding (Long et al., 2005) or through the RHEB promoted activation of phospholipase D1 (PLD1) activation (Sun et al., 2008). TSC1/2 respond to a number of different upstream signals including stress signals such as oxygen levels and energy levels (Hardie, 2007). Reduced energy levels (measured by levels of ATP and ADP) triggers AMPK activation which in turn leads to phosphorylation of TSC2 and subsequent mTOR activation (Inoki et al., 2003). TSC1/2 is the control point of mTOR signaling as crosstalk from other pathways are integrated at this point; for example, TSC1/2 interacts with various pathways including GSK3, AMPK, PTEN/PDK1/AKT (p53/DNA damage repair pathways), IKKβ signaling (inflammatory mediator/cytokine pathways) (Stambolic et al., 2001; Feng et al., 2005; Inoki et al., 2006; Lee et al., 2007). Much of what is known of mTORC1 was elucidated through the usage of a potent inhibitor, rapaymcin. Once inside the cell, rapamycin binds to FK506-binding protein of 12 kDa (FKBP12). This complex then interacts with the rapamycin binding domain (FRB) of mTOR and inhibits mTORC1 function and activities in various processes (Ballou and Lin, 2008).

mTORC1 and translation

Under normal translation conditions, eIF4E binding protein 1 (4E-BP1) is responsible for binding to multiple RNA-binding proteins to prevent the initiation of 5'-cap-dependent translation. These proteins include a number of other eukaryotic initiation factors necessary for stable mRNA translation. In particular, cap-dependent translation is vital for the translation of 5'-untranslated regions (UTRs) that exist on transcripts with high G-C content (usually the mark of extensive secondary structures). In this case eIF4F (one of the initiation factors recruited by eIF4E) is needed for helicase activity to promote stable

binding between the transcript and the ribosome. Thus, alteration of the activity of 4E-BP1 can have downstream effects that may alter the degree to which certain mRNA transcripts are translated (Livingstone et al., 2010). Under conditions favorable for cell growth, mTOR phosphorylates p70 S6 kinase and 4E-BP1 (Beretta et al., 1996; Gingras et al., 1998). Phosphorylation of 4E-BP1 at four sites activates translation by removing 4E-BP1 contact with the translation initiation factors.

Learning, memory, and local translation

Translation in neuronal cells is localized to specific compartments known as RNA granules. This localization is necessary for proteins that need to be expressed in specific compartments or areas of the neuron far from the soma such as axons and dendrites. The production of proteins localized at the synapse are important in the control of the strength and type of signal in synaptic transmission. The consequences of the type of signal and the downstream signaling pathways triggered result in numerous alterations to learning, memory, and the general cognitive landscape (Sutton and Schuman, 2006; Cajigas et al., 2010). The removal of 4EBP2 in the hippocampus results in an increase in excitatory transmission which produces mice with autistic phenotypes (Gkogkas et al., 2013). To this end, alteration of the local translation machinery can have significant effects and dysregulation of normal functioning can lead to disease states.

Long-term memory formation relies on neuronal plasticity to 1) strengthen existing synaptic connections, 2) create new synapses, and 3) regulation of transcription and local translation to modulate ion channel density and conductivity. Thus, long-term memory formation is highly reliant on protein synthesis; to this end, rapamycin treatment has been shown to inhibit long-term memory formation in multiple different contexts (Parsons et al.,

2006; Bekinschtein et al., 2007; Blundell et al., 2008; Gafford et al., 2011; Jobim et al., 2012). S6K1 (vertebrate isoform of S6 kinase, and mTORC1 target) is also important for long-term memory formation as mice with S6K knockouts demonstrate an inability to form memories via early-LTP (Antion et al., 2008b). Additionally, learning processes have been demonstrated to increase S6 activity via mTORC1. These same mice are also impaired in spatial learning. Long-lasting changes to synaptic plasticity seen in memory formation can be induced through N-methyl-D-aspartate glutamate receptor activation (NMDAR). The resulting signaling cascade activates mTOR via the PI3K-mediated route. In this way, NMDAR activation results in the alteration of the protein architecture of the postsynaptic density (PSD) through mTOR-mediated control on local translation in the dendrite.

mTORC1 and the K^+ channel Kv1.1

One specific mechanistic example of the effects of local translation by mTOR can be seen in the regulation of the mRNA transcript for the potassium channel Kv1.1. Kv1.1 is a voltage-gated ion channel that is typically localized in the axons. Kv1.1, like other potassium channels, is responsible for conducting neuronal excitability. Indeed, removal of Kv1.1 in mice results in the development of spontaneous seizures. We previously demonstrated that levels of Kv1.1 at the synapse are altered in an mTOR-dependent manner (Raab-Graham et al., 2006). Treating neurons with rapamycin or inhibiting NMDARs increased dendritic Kv1.1 expression in the hippocampus. We showed that Kv1.1 is localized to the dendrites through a local translation assay that uses Kaede, a photoconvertible fluorescent protein (Chudakov et al., 2010). This suggested that activation of the synapse through NMDARs could have an effect on protein translation in the synapse – in this case a reduction in the synthesis of Kv1.1 in an mTOR-dependent

manner. We identified the mechanism of this reduction by identifying a binding site for the microRNA miR-129 which binds to Kv1.1 mRNA when mTOR is on to repress translation (Sosanya et al., 2013). When mTOR is off, the miR-129 binding site is instead occupied by HuD, an RNA-binding protein that regulates microRNA binding (Kundu et al., 2012). HuD has been shown to have a number of high-affinity targets such as PSD95 and CaMKIIα (Lee et al., 2005; Gong et al., 2006). When these targets are degraded, the free HuD pool increases allowing miR-129 to be displaced by HuD which in turn promotes the active translation of Kv1.1. Herein, we present evidence that Kv1.1 is not the only mTORC1 target that is actively repressed (Chapter 2).

USING HIGH-THROUGHPUT ASSAYS TO UNDERSTAND THE MOLECULAR BASIS OF MEMORY

In order to gain a more in depth view of the underpinnings of synaptic plasticity both in normal and diseased states many laboratories are initiating unbiased screens to identify (1) the mRNA transcripts localized to synaptic compartments, (2) the mRNA transcripts actively translated by the ribosome under specific cellular conditions, and (3) the protein kinases, RBPs, and microRNAs that control the timing and expression of locally translated mRNA. Importantly, researchers are combining classic techniques utilized since the 1950's that have been extended and improved upon with highly specialized high-throughput methods to answer these questions and provide further insights into the molecular basis of neuronal function and neurological disease.

The remainder of this introduction we will focus on the application of these specialized techniques in revealing the complex array of cell signaling and regulatory networks in neurons that govern the processes of local translation and long-term potentiation (LTP) in dendritic spines. First, we will consider experiment design and model

organisms. These techniques have revealed the complex array of cell signaling and regulatory networks that govern local translation and synaptic plasticity in dendrites. We provide a general workflow for large-scale sequencing or proteomics projects highlighting general considerations and caveats at each stage (Figure 1.2). Then, we outline potential methods and strategies to validate findings of these large-scale projects in normal and disease rodent models.

EXPERIMENTAL DESIGN OF HIGH-THROUGHPUT ASSAYS

Large-scale, high-throughput projects that analyze distributions of RNA and protein are generally costly and time-consuming. Thus, careful consideration must be taken at the design stage and each of the following stages to determine the best approach and methodology needed to answer the question of interest (Figure 1.2). After determining any drug or behavioral approaches to the experimental design, the next important consideration is model organism of choice. The majority of research in learning in memory has been dominated by rat, mouse, or human studies but this is by no means the only option (Manger et al., 2008). Typically, rats have been used for cognitive assays because of their large size, low cost, and versatility (Dudchenko, 2004). However, mice have numerous transgenic lines available are easily to manipulate genetically making them a useful model in molecular biology for the study of disease. Furthermore, classic experimental design principles such as randomization are extremely important for high-throughput experiments (Auer and Doerge, 2010; Cui, 2010; Fang and Cui, 2011; Williams et al., 2014). Randomization is vital for lowering noise and ensuring that other unknown variables are not biasing the observed results. The large number of steps involved in high-throughput projects necessitates randomization at every possible step from the design phase to

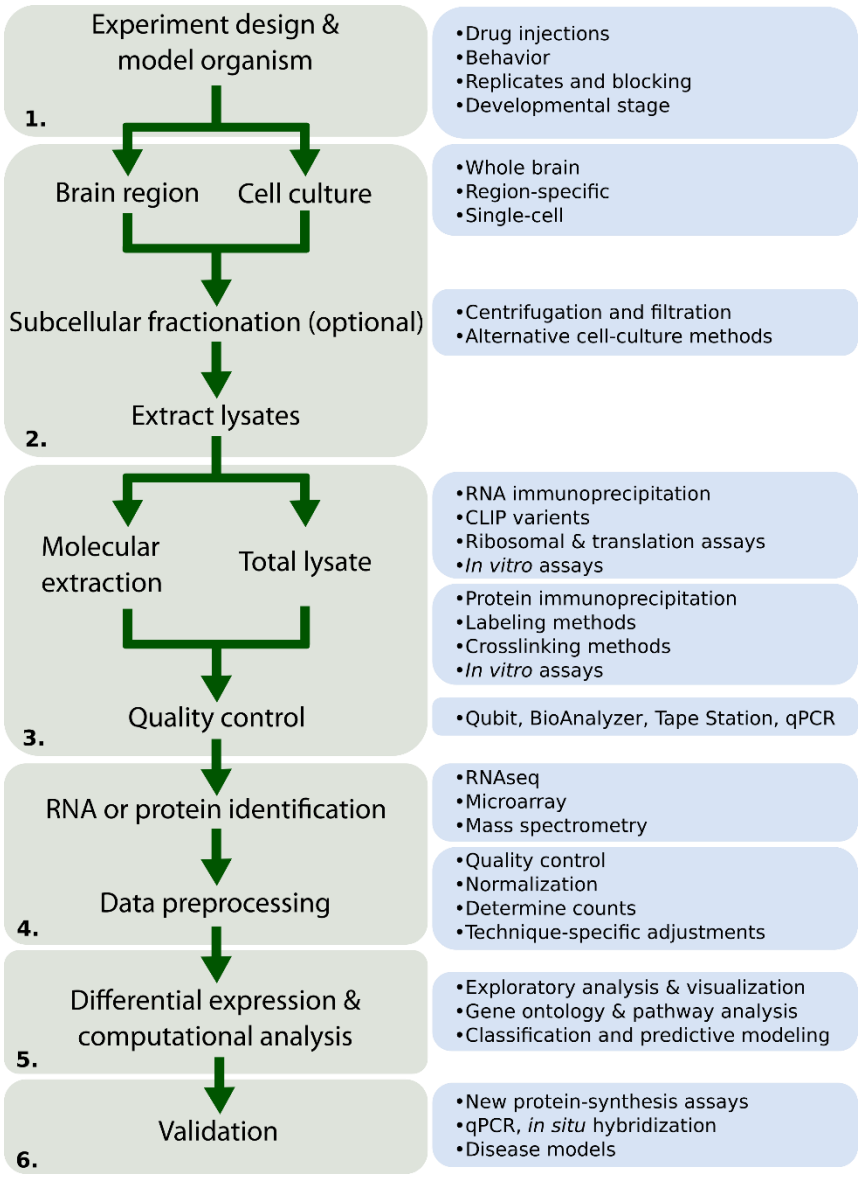


Figure 1.2: Workflow for High-Throughput Projects in the Brain

High-throughput experiments in the brain consist of 6 general stages. **1.** The experiment is carefully designed with attention to replicates, statistical power, cost, and animal model. **2.** Lysates are extracted from a brain region (*in vivo*) or cell culture (*in vitro*) with any subcellular fractionation to study a specific neuronal compartment. **3.** Experiments such as CLIP or Co-IPs are performed to analyze specific interactions between RNA and protein or protein and protein. **4.** RNA or proteins are identified through high-throughput methodology and preprocessed to determine molecular counts. **5.** Analysis is performed on the data to determine new patterns and connections based on the aims of the experiment. **6.** Any conclusions from 5. are validated in animal disease models or through various protein or RNA visualization/quantification assays.

behavior and drug injections. In addition, drug treatments, tissue preparation, and molecular extraction should be performed on the same day for each set of replicates by the same researcher if possible to minimize variability. Replication is also extremely important, particularly for RNAseq experiments and tools such as Scotty (http://scotty.genetics.utah.edu) have been created to aid in the determination of replicate number (Busby et al., 2013; Hart et al., 2013). Due to high variability common in these projects - especially in the case of RNA immunoprecipitations - more biological replicates are extremely important to lowering error. The number of replicates needed for an experiment will vary depending on the type of experiment and ultimately need to be balanced with cost and depth of sequencing but should not be less than three and adding more replicates may gave more statistical power than sequencing deeper (Ching et al., 2014; Liu et al., 2014). To perform a power analysis to determine the number of replicates needed you need a rough estimate of the effect size which in turn depends on the depth of sequencing. Thus, there is a tradeoff between replicates, depth of sequencing, and cost that must be considered during the design phase (Wang et al., 2011; Liu et al., 2013; Vijay et al., 2013).

RATIONALE FOR EXTRACTING CELL-SPECIFIC AND SUBCELLULAR RNA POPULATIONS

Approaches utilized to isolate synaptic mRNAs are vast. Biochemical isolation of synapses via centrifugation or filtration (Figure 1.3) and microdissection of dendritic fields in brain slices have provided a rich source of dendritic/synaptic mRNAs. More recently, the development of technology allowing single-cell RNAseq has allowed researchers to classify cell transcriptome dynamics and determine cell-type diversity (Darmanis et al., 2015; Dueck et al., 2015; Dueck et al., 2016). These single-cell technologies offer

promising opportunities for the field of learning in memory, especially when combined with disease models, electrophysiology, and behavioral studies (Harbom et al., 2016). These data, collectively, will provide powerful models guiding investigators to test translation of specific mRNAs in a cell and site-specific manner.

Utilizing centrifugation, filtration, and density sucrose gradients to isolate dendrites

As early as 1956, researchers have been optimizing biochemical techniques to study synapses in isolation. In 1980, Cheung and colleagues were the first to show that the synaptosome (pre- and postsynaptic nerve endings) fraction contains polyribosomes, making it a promising candidate for the study of synaptic protein synthesis (Verity et al., 1980). The synaptosome (S) preparation combines centrifugation and sucrose gradient fractionation to create a cell fraction containing a sealed presynaptic structure attached to part of the postsynaptic membrane (Hebb and Smallman, 1956; Whittaker et al., 1964). The attached postsynaptic structure varies in size and may even contain the entire (unsealed) dendritic spine. However, the S preparation was seen to be insufficient for properly studying signal transduction and other events that took place in the postsynaptic cell. For this reason, the synaptoneurosome (SN) preparation was developed which includes both the sealed presynaptic structure and sealed postsynaptic compartment isolated through a series of filtration steps and low-speed centrifugation (Hollingsworth et al., 1985). This preparation is now commonly used to study and identify components of the postsynaptic membrane (Figure 1.3). Some of the earliest attempts to characterize proteins of the postsynaptic membranes required methods to further subfractionate the above synaptic fractions (including a slightly modified SN, referred to as synaptodendrosome (SD)) (Rao and Steward, 1991b; 1993). For example, the synaptic

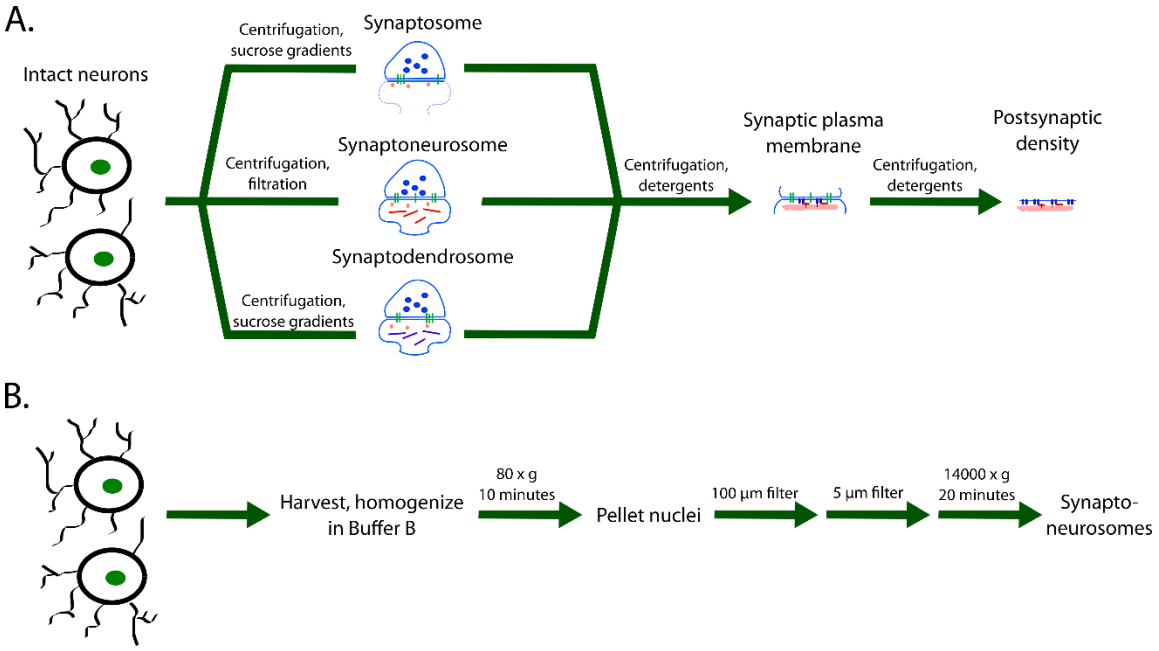


Figure 1.3: Utilizing subcellular fractionation to isolate neuronal compartments

A. Various methods have been developed to isolate neuronal compartments from intact neurons to study their processes in detail. Depicted here are five different subcompartments and their contents. Synaptosomes contain an enclosed presynaptic sac with a portion of the postsynaptic membrane (varies in size), the postsynaptic density, and some of the postsynaptic molecules. Synaptoneurosomes, the most common preparation, consists of an enclosed presynaptic and postsynaptic sac containing many of the molecules involved in synaptic signaling and plasticity. Synaptodendrosomes produce a structure similar to synaptoneurosomes but the content of the posysnaptic sac is not as representative of the intact neuron as the synaptoneurosome. Further centrifugation and application of detergents allows isolation of the synaptic plasma membrane (pre- and postsynaptic membranes with associated ion channels and structural components) and the postsynaptic density. **B.** Typical protocol for extraction synaptoneurosomes modified from Quinlan et al. 1999. Buffer B composition: 20 mM HEPES, pH 7.4, 5 mM EDTA, pH 8.0, protease inhibitor cocktail (Complete, Roche), phosphatase inhibitor, RNase inhibitor Superase-In at 40 units/ml (Ambion).

plasma membrane (SPM) subfraction, containing the proteins embedded in the plasma membrane around the synapse, can be purified from the synaptic fractions (Blackstone et al., 1992; Bermejo et al., 2014). The SPM fraction can be further fractionated to obtain the PSD subfraction, the large scaffolded complex of proteins found clustered at the edge of the postsynaptic membrane (Carlin et al., 1980; Villasana et al., 2006). Rao and Steward isolated SPMs by subjecting S to further gradient fractionation combined with the detergent Triton X-100 and demonstrated that protein synthesis was occurring in this region (Rao and Steward, 1991a). Expanding on this early finding by Rao and Steward, Niere et al. subjected the PSD fraction, isolated in a similar manner, to mass spectrometry to show that 75% of the PSD changes in composition upon inhibiting the protein synthesis pathway mTOR for only one hour *in vivo* (Niere et al., 2016). Thus, combining classic biochemical subcellular techniques with big data approaches has vastly expanded our knowledge in how dynamic protein synthesis occurs in the PSD.

Limitations to biochemical synapse isolation and potential strategic measures utilized to overcome these limitations.

Though subcellular fractionation is still commonly employed to study molecular events at the postsynapse, there are limitations. The S, SN, and SD preparations are considered impure since they cannot successfully remove glial fragments (Chicurel et al., 1990; Rao and Steward, 1991b; 1993). Hollingsworth and colleagues noted "unidentifiable debris" when they analyzed the preparation by EM which suggests that there may be foreign protein and RNA carryover from membrane fragments or other fractions not normally associated with axons or dendrites (Hollingsworth et al., 1985). An additional concern is the presence of somatic contaminants. EM provides the most rigorous check, however one might quickly screen for somatic contamination by examining a sample of

the synaptic preparation versus the total homogenate with a nuclear stain such as DAPI or western blotting for a nuclear protein such as NeuN (Sosanya et al., 2013). Moreover, after large scale screens one may subtract possible contaminants bioinfomatically. To determine axonal or dendritically expressed transcripts, Schuman and colleagues sequenced RNA isolated from the stratum radiatum and lacunosum moleculare of the rat hippocampus and subtracted transcripts enriched in glia, interneurons, nucleus, mitochondria, and blood vessels based on cell type-specific transcriptome data from previous publications and online databases (Cajigas et al., 2012). While this technique may eliminate candidates that are expressed both in glia and neurons, it does allow one to follow up putative dendritic mRNAs with more certainty. In spite of these limitations, these preparations have provided the basis for many studies that have moved the field forward.

Alternative approaches to isolating dendrites

Eberwine and colleagues were the first to isolate dendrites from neuronal cell bodies in hippocampal cell cultures. A micropipette was used to microdissect the dendrite in which RNA was isolated. Differential display and microarray analysis of mRNA isolated in this manner, provided the first large-scale analyses of the mRNA present in the dendrite (Miyashiro et al., 1994; Eberwine et al., 2001; Miyashiro et al., 2003). Considering the limitations of these early assays, remarkably the authors estimated that ~400 mRNAs reside in the dendrites (Eberwine et al., 2001).

Since then, less labor-intensive methods have been developed. One clever technique capitalized on the fact that neuronal and glial cell bodies are typically at least 10 µm in size. By plating hippocampal neurons onto PET membranes with 3 µm pores neuronal processes are separated from neuronal cell bodies and glia. Thus, neuronal

processes can be isolated by scrapping the bottom of the filter (Torre and Steward, 1992; Poon et al., 2006). Additionally, laser capture microdissection of neurites has also been employed successfully to catalog the mRNAs present in dendrites (Kye et al., 2007). More recently, Eberwine and colleagues have refined this method to isolate a single neuron from cell culture including dendrites (Lovatt et al., 2015). These techniques offer promising alternatives to complement the standard subcellular fractionation methodologies. Through mRNA amplification technologies, it is now possible to perform cell and compartment-specific identification of synaptic mRNAs.

WHAT ISOLATING REGULATORY FACTORS CAN TELL YOU ABOUT SYNAPTIC EFFICACY

Isolation of RNA-Binding proteins (RBPs) and RNA populations

RNA-binding proteins (RBPs) and all types of RNAs are believed to play a vital role in learning and memory by controlling transcript localization and availability at the dendrite (Sephton et al., 2011; Aksoy-Aksel et al., 2014; Lenzken et al., 2014; Smalheiser, 2014; Zhou et al., 2014). In the past decade, a number of important high-throughput techniques have been developed concurrently with specialized deep sequencing technology that has allowed researchers to elucidate the RNA populations bound to RBPs or ribosomes on an unprecedented scale. However, these techniques are very difficult to perform and must be optimized considerably. In the following sections, we will outline the basics principles of these techniques and compare their advantages and disadvantages (Table 1.1). We will also consider a number of *in vitro* selection-based techniques that are easier to perform and may reveal information to complement other assays.

RNA immunoprecipitation sequencing/microarray (RIP-SEQ/RIP-CHIP)

The RNA immunoprecipitation (RIP) has been used previously to identify targets of RBPs involved in neurological dysfunction (Buckanovich and Darnell, 1997; Napoli et al., 2008; van der Brug et al., 2008; Fernandez et al., 2015). High-throughput RIP-SEQ serves as a useful tool in determining RNA populations bound to proteins involved in local translation. RIP, similar to the protein-based immunoprecipitation procedure, has been optimized in order to preserve the RNA-protein complex during the lysing step such as gentle-freeze thawing (Keene et al., 2006; Oeffinger et al., 2007; Jain et al., 2011; Dahm et al., 2012). Other modifications in RIP protocols are to ensure that free RNAs released during lysis do not bind non-specifically to the beads or the RBP, a phenomenon that has been observed previously and contributes to the high background of some RIP experiments (Mili and Steitz, 2004). Following some of the procedures outlined by Jain and colleagues, it is possible to achieve minimal or negligible levels of background binding (Jain et al., 2011). While crosslinking with formaldehyde to bind the protein-protein or protein-RNA structures together may help in ensuring complex isolation, in some cases it may not lower the level of background binding (Penalva et al., 2004). Finally, like many other high-throughput techniques, there are limitations including epitope accessibility to the antibody, as well nonspecific binding inherent with antibody-based procedures. Importantly, RIP is not able to reveal precise binding sites like other crosslinking techniques discussed below; however, it can reveal the full-length transcript of RNAs bound to the protein of interest *in vivo*.

Assay Type	Technique	Advantages	Disadvantages	Reference
Immunoprecipitation	RIP-SEQ/RIP-CHIP	Recover full length-RNA	High background, antibody-based	(Keene et al., 2006; Oeffinger et al., 2007; Jain et al., 2011; Dahm et al., 2012)
RNA-Protein interactions through CLIP-SEQ	HITS-CLIP	High resolution	Difficult, low cross-linking efficiency, cross-linking artifacts, antibody-based, RT-PCR mispriming, cannot distinguish between single protein binding and protein complex binding	(Licatalosi et al., 2008; Darnell, 2010; Kishore et al., 2011; Zhang and Darnell, 2011; Gillen et al., 2016)
	PAR-CLIP	Very high resolution, high cross-linking efficiency	Difficult, expensive, 4-SU toxic, high background, antibody-based, low alignment %	(Hafner et al., 2010; Spitzer et al., 2014)
	iCLIP/iCLAP	Very high resolution, RT-PCR does not stall at crosslink site	Very difficult to perform	(Konig et al., 2010; Wang et al., 2010b; Sugimoto et al., 2012; Huppertz et al., 2014)
	CRAC	Affinity purification-based, less background	Difficult, tag may interfere with protein function	(Granneman et al., 2009; Bohnsack et al., 2012)
	PIP-SEQ	Does not use UV cross-linking, identifies non-Poly(A) transcripts	Difficult, very new method	(Silverman et al., 2014)
RNA Structure	CLASH-SEQ	Identification of RNA-RNA duplexes	Use of two adaptors results in ambiguity, ligation reaction inefficient	(Kudla et al., 2011)
	HiCLIP	Improves upon CLASH-SEQ, can identify long RNAs	Difficult, very new method	(Sugimoto et al., 2015)
Ribosome-based	RIBO-SEQ	Greatly improves on past footprinting techniques, high-throughput	Lysis preparation may change ribosomal distribution, stalled ribosome may bias results	(Esposito et al., 2010; Masek et al., 2011; Gandin et al., 2014)
	TRAP-SEQ	Easier to perform than RIBO-SEQ	Lacks RIBO-SEQ specificity	(Jiao and Meyerowitz, 2010)
In vitro binding	SELEX	Quick, easier to perform than alternative *in vivo* methods	High affinity motif bias, identifies non-physiological interactions	(Darmostuk et al., 2015)
	RNAcompete	Quick, easier to perform than alternative *in vivo* methods	RNA secondary structures may affect binding assay, identifies non-physiological interactions	(Ray et al., 2009)
	SEQRS	Identifies many more motifs than RNAcompete	Identifies non-physiological interactions	(Campbell et al., 2012)
	RNA Bind-n-Seq	Greatly improves on other *in vitro* methods, compliments CLIP-based assays	Identifies non-physiological interactions	(Lambert et al., 2014)

Table 1.1: Comparison of techniques to isolate RNA-associated populations

Crosslinking-based techniques to identify RBP-bound RNAs

HITS-CLIP

Cross-linking and immunoprecipitation (CLIP) was originally developed by the Darnell Lab to study interactions between the neuronal protein Nova and its target RNAs (Ule et al., 2003; Ule et al., 2005). When combined with high-throughput sequencing, the modified protocol is referred to as HITS-CLIP (Licatalosi et al., 2008; Darnell, 2010). The HITS-CLIP technique allows researchers to perform mapping of RBP binding sites on RNA in a high-throughput manner. The procedure has now been expanded and modified to a number of variants which we will detail below. Modifications have also been made to the HITS-CLIP procedure that now allow up to single-nucleotide resolution of RBP binding sites which could aid in the discovery of RNA-binding motifs utilized in local translation (Kishore et al., 2011; Zhang and Darnell, 2011). While HITS-CLIP allows for fine resolution of RNA-protein interaction sites, the crosslinking procedure can introduce artifacts and during reverse transcription mispriming events can occur (Kishore et al., 2011). Notably, recent improvements have been made to the procedure to minimize mispriming artifacts through the use of two specialized primers during the reverse transcription step (Gillen et al., 2016). One of the biggest limitations to HITS-CLIP is the low cross-linking efficiency, which has been reported to be ~5% (Darnell, 2010). HITS-CLIP and its derivatives have already been utilized to identify RBP binding sites for proteins involved in local translation, to determine microRNA bindings sites, and identify RNA targets for proteins involved in neurological and developmental dysfunction (van der Brug et al., 2008; Darnell et al., 2011; Ascano et al., 2012b; Ince-Dunn et al., 2012; Lagier-Tourenne et al., 2012; Wagnon et al., 2012; Boudreau et al., 2014; Weyn-Vanhentenryck et al., 2014; Scheckel et al., 2016).

The first step in the HITS-CLIP procedure is to use UV light is used to crosslink RNA and protein interacting within the cell population of interest. Next, cells are lysed and an RNase is used to shorten the crosslinked RNA-protein complexes in the cell lysate. The complexes are then immunoprecipitated and a radiolabeled adaptor (a nucleotide sequence used in high-throughput sequencing) is added to the 3' end. After elution from beads the protein-RNA complex is isolated by SDS-PAGE and transferred to a nitrocellulose membrane. The RNA of interest is then isolated from the membrane by proteinase K treatment and an adapter is added to the 5' end. The isolated RNA can then be prepared for high-throughput sequencing. HITS-CLIP allows for fine resolution of RNA-protein interaction sites however, the crosslinking procedure can introduce artifacts and mispriming events can occur during reverse transcription (Kishore et al., 2011). Recently, improvements have been made to the HITS-CLIP procedure to address mispriming artifacts (Gillen et al., 2016). HITS-CLIP also suffers from low cross-linking efficiency - a maximum of 5% (Darnell, 2010).

PAR-CLIP and iPAR-CLIP

Photoactivatable-ribonucleoside-enchanced crosslinking and immuniprecipitation (PAR-CLIP) was introduced in 2010 to address issues of low crosslinking efficiency in HITS-CLIP and issues of noncrosslinked background RNA in samples (Hafner et al., 2010; Spitzer et al., 2014). A photoreactive nucleoside analog of uridine (4-SU) and guanosine (6-SG) are added to cultured cells which increases crosslinking efficiency. 4-SU causes a thymidine to cytidine transition during the reverse transcriptase reaction thus indicating the exact crosslink sites. 4-SU was not originally believed to be toxic to cells at concentrations used in PAR-CLIP but recent evidence suggests it may inhibit processing of 47S rRNA and affect the experimental results (Burger et al., 2013). PAR-CLIP has been limited to use

in cell culture as its toxicity in animal tissue is not yet known but an *in vivo* approach called iPAR-CLIP has been developed, with the "i" standing for *in vivo* (Jungkamp et al., 2011). Furthermore, PAR-CLIP was quantitatively shown to have reproducible levels of background signals thus necessitating the use of empirically determination of background as an extra step in analysis (Friedersdorf and Keene, 2014). PAR-CLIP also suffers from other issues as aligned reads can be as low as 20% of total reads (Hafner et al., 2012). Importantly, PAR-CLIP has been used to understand the binding of RBPs whose dysregulation has been shown to play a role in neuronal diseases (Ascano et al., 2012a). Some example incude Rbfox3 which was find to have a unique function in the regulation of pri-mRNA (Kim et al., 2014), the first identification of two FMRP binding motif sequences (Ascano et al., 2012b), and the first report to identify all the targets of the FET protein family (FUS, EWSR1, and TAF15) (Hoell et al., 2011). Thus, PAR-CLIP has moved the field forward by determining binding motifs/targets for RBPs allowing investigators to answer question on how RBPs contribute to coordinated translation with synaptic plasticity.

iCLIP, iCLAP, and CRAC

Individual-nucleotide resolution CLIP (iCLIP) was developed in response to the data showing that the reverse transcription reactions truncate at the crosslink sites in HITS-CLIP and PAR-CLIP (Konig et al., 2010; Sugimoto et al., 2012; Huppertz et al., 2014). The iCLIP method adds a circular PCR amplification step that allows researchers to determine the sequence of cDNAs that would normally be truncated in other CLIP methods. The method is similar to HITS-CLIP in execution. Immunoprecipitated RNA-protein complexes are treated with proteinase K; crosslinked proteins will not be affected and stick to the binding site. During cDNA synthesis, truncation will occur at the crosslink

site. The cDNA is then circularized, linearized, and PCR amplified to determine the region of the protein binding site at the crosslink. While iCLIP does resolve some of the issues that other CLIP methods face it is technically challenging and has extra steps to handle RNA which is already in limited quantities. Additionally, the extra manipulation at the PCR amplification stage could bias the final results. iCLIP has another variation known as Individual-nucleotide resolution crosslinking and affinity purification (iCLAP) which uses a two-step affinity purification. This technique may be an option if antibodies are not available and may lower background (Wang et al., 2010b). Finally, another technique involving affinity purification is cross-linking and analysis of cDNAs (CRAC). This technique requires RBPs to be tagged with a protein A and hexahistidine for IgG purification followed by nickel-affinity purification (Granneman et al., 2009; Bohnsack et al., 2012). This technique has been used to uncover spliceosomal RNA-protein interactions and may prove to be a useful method for cleanly isolating a protein within a complex. Thus, each technique has been optimized to overcome specific limitation, to provide researcher with a tool kit to address their specific question (Table 1.1).

PIP-SEQ, HiCLIP, and CLASH-SEQ

Protein interaction profile sequencing (PIP-SEQ) is another more recent high-throughput method that can map RNA-protein interactions in an unbiased, transcriptome-wide manner, rather than selectively with specific RBPs (Silverman et al., 2014). Furthermore, RNA secondary structures are an important regulatory mechanism in post-translational control. Indeed, RNA secondary structure has already been observed as a control mechanism in long-term memory formation as well as a 3'UTR recognition sequence for localization to the dendrite (Martin and Ephrussi, 2009). HiCLIP and CLASH-SEQ are two related methods that can be used to map RNA secondary structures

(Kudla et al., 2011). CLASH-SEQ is a high-throughput method that allows for the transcriptome-wide level identification of secondary structures via analysis of RNA duplexes. HiCLIP improves upon the biases and limitations of CLASH-SEQ by adding another adapter that allows identification of RNA-RNA duplexes with greater precision (Sugimoto et al., 2015). Collectively, these techniques are providing answer to long-sought after question regarding how secondary structure may ecncode dendritic targeting and translational regulation signal that investigators have struggled with for several years.

RNA Interactome Capture

Another crosslinking method that differs from the CLIP-based methods has recently been developed. RNA interactome capture can be used to survey the full repertoire of both protein and RNA interacting physiologically within cells (Castello et al., 2013; 2016). UV irradiation is used to crosslink RBPs to polyadenylated RNAs which are then isolated using oligo(dT) magnetic beads. Next, RNA and protein are separated and analyzed by RNAseq and mass spectrometry respectively. Like all crosslinking methods, it is limited by crosslinking efficiency. Furthermore, it will not be able to isolate RBPs bound to nonpolyadenylated RNA. Thus, for the first time, investigators can isolate protein-RNA interactions as a network, providing insight into how RBPs work in concert to regulate mRNA translation of plasticity-related proteins.

Ribosomal/Translation-Based methods

RIBO-SEQ/ARTSEQ and Polysome Profiling

The analysis of global mRNA levels within a cell population is commonly used to measure gene expression. However, this may not be a sufficient metric as mRNA levels do

not necessarily correlate to protein expression levels due to an extra layer of translational control at the level of the ribosome. Therefore, ribosome-specific RNA methods have been developed to better understand the dynamics and control of mRNA translation. Translation serves as a rapid mechanism by which the cell can finely control the amount of protein to be expressed from a particular mRNA in both the spatial and temporal dimensions. Such regulation of translation serves a major function in both memory formation and synaptic plasticity thus necessitating the need for methods able to profile mRNAs under active translation (Costa-Mattioli et al., 2009; Buffington et al., 2014). One approach to identifying mRNAs under active translation is polysome-profiling in which ribosomes with high translation efficiency are selectively isolated by polysome gradient fractionation, followed by RNA isolation and high-throughput sequencing or microarray (Esposito et al., 2010; Masek et al., 2011; Gandin et al., 2014). Ribosome profiling sequencing (RIBO-SEQ or active mRNA translation sequencing, ARTSEQ) is a genome-wide approach used to identify mRNA being actively translated by the ribosome without consideration of translational efficiency. In this context, translational efficiency is defined as the mean ribosomal footprint counts for a given mRNA, a quantitative measure of the degree of ribosomal occupancy (Ingolia et al., 2009; Ingolia et al., 2012). All ribosomes in active translation are isolated and the associated untranslated mRNA is then removed and digested. Then, the rRNA is depleted from the samples and the actively translated mRNA is reverse transcribed and sequenced. The fragments of RNA protected from digestion are then mapped to a reference genome thereby providing the location of the ribosome on various mRNA at nucleotide-scale. Since ribosome profiling uses a footprinting approach, it is able to reveal the precise binding sites of the ribosomes across the mRNA and provide quantitative measures of expression. These features are not possible with the traditional polysome profiling approach. Due to the high sensitivity of RNA sequencing approaches,

these methods can provide detailed information about mRNA undergoing translation, thus allowing researchers to better understand how the synapse is actively changing in the context of learning and memory.

TRAP/TRAP-SEQ and RiboTag

Translating ribosome affinity purification (TRAP) is another method for mapping actively translated mRNAs using EGFP-tagged ribosomal protein L10 (RPL10). The technique is performed in transgenic mice containing Bacterial Artificial Chromosomes (BACs) (Heiman et al., 2008). The technique was later extended for use in RNAseq in a method known as TRAP sequencing (TRAP-SEQ) which uses His and FLAG epitope-tagged ribosomal protein L18 (RPL18) to immunopurify translating ribosomes (Jiao and Meyerowitz, 2010). However, because this assay captures whole ribosomes (both polysomes and monosomes) the translational state of the mRNA of interest will not be as high-resolution as that obtained from ribosomal profiling or polysome profiling. A recent study has attempted to modify the TRAP-SEQ method in such a way that it is possible to extract ribosome-bound mRNA specifically from dendrites (Ainsley et al., 2014). RiboTag is another recently developed method for a mouse transgenic line in which the ribosomal protein L22 (RPL22) gene has been tagged with an HA tag before the stop codon. This mouse can then be crossed with a mouse line contained cell type-specific Cre-recombinase thus creating HA-tagged ribosomes in the cell-type of choice. Immunoprecipitation will recover ribosome-bound mRNA in the chosen cell type (Sanz et al., 2009). This method circumvents the need for the BAC required in TRAP-SEQ. Similar to the HA-tagging and TRAP technology described above, another method has been recently developed that allows ribosomes to be GFP-tagged, but only immunoprecipitated from cells that project to a specified brain region (Ekstrand et al., 2014). Thus, TRAP-SEQ provides extensive

information about mRNA populations undergoing translation and with the modification of circuit-specific GFP-tagged ribosomes it's now possible to examine coordinated mRNA translation between specific brain regions.

In vitro binding assays

RNA-SELEX, RNAcompete, SEQRS, and RNA Bind-n-Seq

In vitro binding assays provide a means for surveying the RNA-binding preference of RBPs. While some of the results may be non-physiological, they are useful for motif identification and can complement other antibody-based methods as a means separating falsely-identified RNA targets. Systematic evolution of Ligands by EXponentional Enrichment (SELEX) was developed in the 1990s as a way of assessing the binding affinity of proteins to a pool of random oligonucleotides. The oligonucleotide library is incubated with the target protein of interest. Candidate oligonucleotides that bind the protein are reverse-transcribed, amplified, and used to seed a new round of selection with the protein of interest. After several rounds, the RNA is sequenced. These RNA molecules represent sequences with high-affinity to the protein of interest. SELEX is useful for determing novel RNA-protein interactions as well RBP motif discovery. Many improvements have been made to SELEX over the years and RNA SELEX now exists in a high-throughput form, giving researchers the ability to assess possible RNA sequences that can bind to a given molecular target (Darmostuk et al., 2015). Other attempts to assess RBP binding preference include RNAcompete (Ray et al., 2009). Here, a custom-made microarray is used to produce a pool of RNAs (29-38 nucleotides in length) which are either unstructured or contain stem-loops. These RNA molecules are then made double-stranded through primer extension on the array. After release from the array, GST-tagged RBPs are incubated with

the RNA pool. The RNA bound to the RBP is then removed, extracted, labeled, and hybridized to a microarray for high-throughput analysis. SEQRS is another method that builds upon older forms of *in vitro* selection such as RNA SELEX. DNA oligonucleotides with a random 20 nucleotide sequence are transcribed to RNA and then incubated with a recombinant protein of interest. The RNAs are then extracted, converted to cDNA, and sequenced (Campbell et al., 2012). RNA Bind-n-Seq was recently developed with an aim to improve upon the above methods as well as provide a method that was not as challenging, time-consuming, or biased as CLIP (Lambert et al., 2014). The method used a random pool of RNAs which are incubated with a purified RBP present at different concentrations. The RBPs are pulled-down using streptavidin magnetic beads, the RNA is extracted, converted to cDNA, and sequenced. The authors suggest that this method be used in tandem with CLIP-based techniques to filter out false positives.

HIGH-THROUGHPUT ASSAYS

Following isolation of the samples of interest, the samples must be processed and then submitted for high-throughput analysis. For RNA samples, the newest technology available is generally referred to as RNAseq. In the last decade, RNAseq has replaced microarray technology primarily because it is believed to be more accurate, more sensitive, and has a broader dynamic range (Marioni et al., 2008; Fu et al., 2009; Zhou et al., 2014). However, RNAseq is still a very new technology. For this reason, it is more expensive and the analysis pipelines are not completely standardized. RNAseq pipelines also require in-depth bioinformatics analysis though many sequencing facilities offer analysis services. Like microarrays, mass spectrometry is an older technique with a more standard analysis

pipeline that has been reviewed in detail elsewhere (Slonim and Yanai, 2009; Lavallee-Adam et al., 2015).

RNA-Sequencing (RNAseq)

Next-generation sequencing technology (NGS) represents the latest technologies used in high-throughput sequencing. A number of different platforms are available including Illumina, Ion Torrent (Fischer), Roche 454 (Roche), and SOLiD (Life Technologies). Various approaches to sequencing have been developed and improvements have been made over the years (Goodwin et al., 2016). Here we will focus on solid-phase bridge amplification technology pioneered by Illumina. The Illumina platform is currently the most widely used and the company's HiSeq 2000 boasts the lowest sequencing cost per Gb and a low error rate (Loman et al., 2012; Quail et al., 2012). However, for greater sensitivity and a lower error rate (which may be of use for SNP analysis, for example), the RSII platform (Pacific Biosciences) may be preferred. The entire RNAseq pipeline is summarized in Figure 1.4.

Quality control

Before sequencing begins, RNA obtained from cell samples are used to generate a cDNA library (Head et al., 2014). A common pitfall is degraded or impure RNA. Low quality RNA generates very noisy data whose sequences can be difficult or impossible to reconstruct during the data preprocessing stages. Furthermore, sequencing degraded RNA leads to high variability and can impact interpretation and differential gene expression (DGE) analysis leading to the possibility of overfitting in classification analysis downstream. There are a few technologies available for RNA quality and quantity

assessment prior to library preparation. *Measuring quantity:* A UV light-based Nanodrop (Thermo Fischer Inc.) may be used but it is not as sensitive as dye-based fluorescence methods of nucleic acid detection, particularly in the low range. The Qubit 3.0 Fluorometer (Thermo Fischer Inc.) provides accurate quantity assessment at low concentrations (as low as 250 pg/uL using the RNA high-sensitivity kit), even if the RNA is degraded, and is the preferred instrument for quick quantitative assessment. *Measuring quality*: The RNA integrity number (RIN) is a statistical measure of RNA integrity that has been developed to assess RNA quality in a particular sample (Schroeder et al., 2006). RIN values range from 1-10 and typically a RIN of 7-10 is recommended for library construction unless the RNA was derived from rare tissue and obtaining higher quality RNA is costly or impossible. However, it appears possible to construct libraries and analyze data from moderately degraded RNA (RIN = 4-6) using appropriate statistical corrections (Gallego Romero et al., 2014; Sigurgeirsson et al., 2014; Cieslik et al., 2015). The Agilent Bioanalyzer 2100 (Agilent Technologies) and the Agilent Tape Station (Agilent Technologies) both utilize microcapillary based electrophoresis to analyze RNA quantity as well report an accompanying RIN value. As an alternative to using RIN, qPCR may also be used for quality assessment if potential transcripts are already known (Vermeulen et al., 2011). For RNA-protein interaction studies, qPCR can used to assess the validity of target samples by using primers for known RNA targets of the immunoprecipitated RNP. This method does not assess the entire extracted RNA sample but can be a useful tool to demonstrate the success or relative quality of the RNA isolation assay.

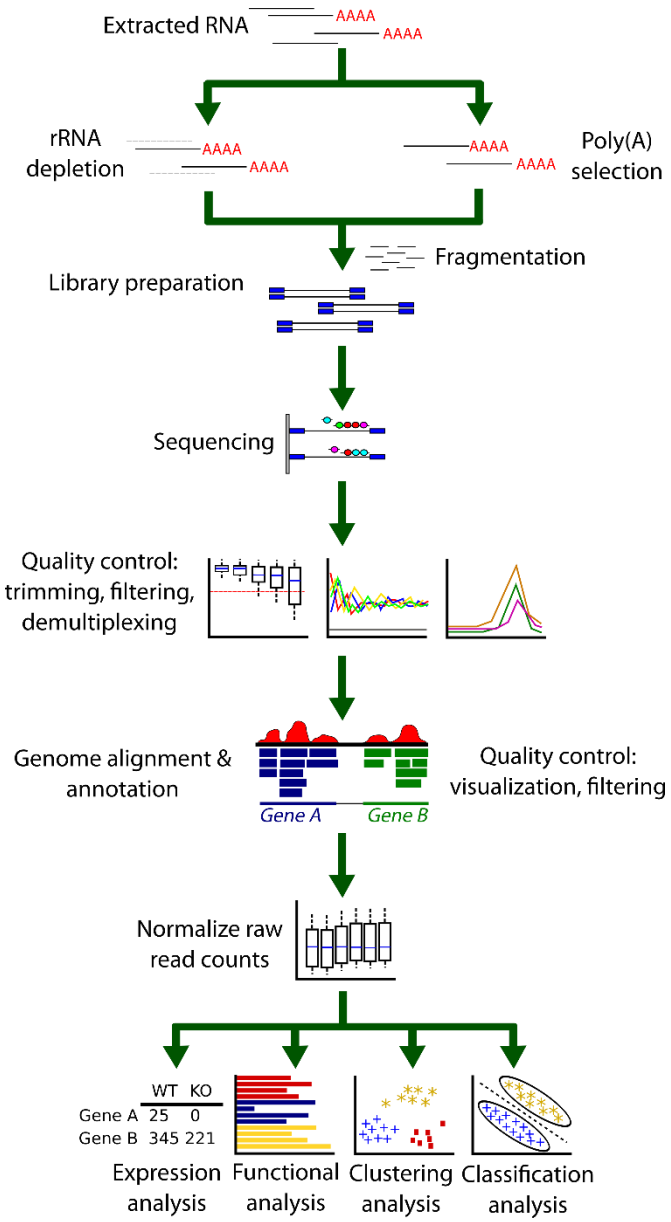


Figure 1.4: General RNAseq pipeline for differential gene expression analysis

Extracted mRNA typically undergoes rRNA depletion or Poly(A) selection to remove the rRNA population from the sample. After fragmentation and library preparation, the sequencing begins. After sequencing, the reads must be checked for quality control and low-quality reads are trimmed or discarded. Next, the reads are aligned to a genome and successfully annotated libraries can undergo visualization and further quality control. Reads are then counted for each gene and normalized so they can be compared. Finally, the gene-reads can be analyzed by a number of methods to discovery new interactions and patterns in the data.

rRNA removal

After RNA quality control, the samples typically undergo a selection stage for the removal of ribosomal RNA (rRNA). Since RNAseq will provide sequences for the most abundant RNAs in the sample, rRNAs must be removed since they comprise > 90% of the RNA isolated from a cell population. There are two major strategies to approach this issue and recent work suggest that both methods introduce biases in sequence coverage (Lahens et al., 2014). However, both of these methods are currently in wide use and since no alternatives exist, efforts are ongoing to decrease bias (van Dijk et al., 2014). Since mature, processed mRNAs contain a poly(A) sequence, oligo(dT) beads can be used to select for only mature mRNAs in a technique called poly(A) selection, effectively removing the rRNA from the sample. While cheaper and of higher sensitivity compared to the alternative choice, the rRNA depletion technique, it will not pick out non-coding RNAs and may detect less genes overall (Cui et al., 2010; Zhao et al., 2014). There are still concerns of 3'end bias for poly(A) selection sequences but recent studies suggest that this may be overcome with appropriate statistical corrections (Roberts et al., 2011). rRNA depletion utilizes beads consisting of sequences complimentary to rRNA. When RNA samples are introduced to the beads, the rRNA attaches to the beads allowing the remaining RNAs to be recovered. Currently, this method is more expensive and does not remove all rRNA (Cui et al., 2010). However, it is necessary for sequencing non-coding RNA.

High throughput sequencing

Following rRNA depletion, the RNA is sheared into small fragments and converted to cDNA by reverse transcription. Unique DNA linkers are ligated onto the 3' and 5' ends of the isolated oligonucleotide. Special capture sites are then added on to the 3' and 5' ends of the molecule which allows them to be anchored onto a solid support surface during the

sequencing steps. PCR amplification is often performed at this stage depending on the amount of RNA available. RNA sequencing is often multiplexed so multiple experiments can be run simultaneously. This is accomplished by attaching a barcode between the 3' linker and the 3'end capture site. These unique barcodes can be used to differentiate between multiple sets of samples run simultaneously (either one barcode, single index, or two, dual index). The samples are then loaded into the sequencer where they are washed over the surface of a small, thin surface known as a flow cell and separated into single-stranded DNA (ssDNA). The samples attach to the flow cell via the ligated capture sites. Since there are capture sites on both sides, the molecule can bend over to attach to the flow cell at both ends. A polymerase and dNTPs are introduced and to amplify the DNA which is separated again into ssDNA. This process is repeated several times within an area of the flow cell to form a cluster of replicated molecules. The flow cell is then exposed to a polymerase and fluorescent dNTPs. The dNTPs are added one at a time to a cluster which fluoresces a different color depending on the incorporated base until the full sequence has been determined. The Illumina technology allows sequencing from just one side (single-read, SR) or from both ends (paired-end, PE). Paired-end runs, which increase the accuracy of genome mapping, are typically more expensive and more useful for difficult, repetitive genomic sequences or providing information about spice junctions and alternative splicing transcripts (Williams et al., 2014). PE is not required for differential gene expression analysis but may provide greater coverage during the genomic alignment stage. Each machine offers a maximum achievable number of output sequence reads per lane of a flow cell (with a total of eight lanes per flow cell). Therefore, the number of samples loaded per lane must be divided by the total number of possible reads to determine the number of reads each sample may receive (the HiSeq 4000 typically delivers 300-400 million total reads per lane). A standardized number of reads required per sample to successfully map

to the genome, determine gene expression differences, or other experimental parameters has not yet been determined and must be considered on an experiment-to-experiment basis. Though the ENCODE consortium has released guidelines for standardized RNAseq practices they are now out-of-date given the pace the field is moving. However, some general guidelines and standards are available (Williams et al., 2014; Conesa et al., 2016).

Data preparation pipeline

There is no one standard analysis pipeline for RNAseq projects as they will vary given the data. Here we present a basic overview of some considerations and software for the DGE analysis process and other downstream analysis (Oshlack et al., 2010; Rapaport et al., 2013; Williams et al., 2014; Finotello and Di Camillo, 2015; Conesa et al., 2016). Also, note that CLIP studies may require specialized experimental design and analysis using specific software (Wang et al., 2015; Zhang et al., 2015). A selected list of currently popular tools for RNAseq data processing and analysis is shown in Table 1.2.

The first major consideration before beginning RNAseq analysis is that the process requires a considerable amount of computing power. The amount of power needed will vary depending on the size and type of experiment. As a general minimum guideline: Large disc space (1-5 TB, solid-state preferred), RAM (at least 8-32+ GB, higher is preferred), and a multi-core CPU (4 preferred, and 2.6+ GHz clock speed) are essential for fast computation; the processing can easily take days on a slower machine and will thus be unavailable for any other use during this time. GPU computing has not currently been optimized for RNAseq. Some desktop computers and laptop builds may be suitable for small-scale analysis but typically workstations or specialized computational facilities (available at many universities) are preferred. Analysis may be performed using

R/Bioconductor packages (https://www.bioconductor.org), specialized software with graphical interfaces (e.g. Galaxy - (Afgan et al., 2016)), programming and scripting languages (e.g. R, Matlab, Python, Ruby, Perl, Java etc.), and terminal commands on a UNIX-based operating system. Analysis and preprocessing is possible on Windows operating systems but may be difficult to set up because a number of bioinformatics tools have been developed for a terminal/command line interface that is not normally compatible with the Windows environment. A Linux setup is the most convenient for analysis but it is possible to set up some tools on Windows usage the Cygwin terminal.

RNAseq preprocessing and analysis roughly falls into the following pipeline:

(i) *Process reads obtained from sequencer*: The sequencer will return the sequencing data in the form of FASTA files. Illumina sequencers perform their own QC checks by default but to determine if the sequences need further processing, the files should be analyzed by a QC tool such as FastQC (Andrews, 2010), ShortRead (Morgan et al., 2009), or RNAseq specific tools RSeQC (DeLuca et al., 2012; Wang et al., 2012a) and RNA-SeQC. These tools return a series of graphs and metrics that can be used to evaluate whether or not any data cleaning is necessary. Each FASTA file lists sequences accompanied by a base quality score (Q score) indicating the probability of an incorrect base assignment for that position in the sequence. The average Q value can be detected using a QC tool. A good sequencing read will have a mean Q value over 30. However, many sequences will just have a drop in Q value toward the 3' end of the sequence and may need trimming by a few bases if they are extremely poor quality. Although trimming reads might be necessary in some special cases, there has been recent work to suggest that overtrimming can affect differential gene expression estimates at later analysis stages (Williams et al., 2016). Finally, the adapter sequence might need to be removed if it was detected

Tool Type	Tool Name	Reference
Quality Control	FastQC	(Andrews, 2010)
	RNA-SeQC	(DeLuca et al., 2012)
	RSeQC	(Wang et al., 2012a)
	ShortRead	(Morgan et al., 2009)
Trimming, Demultiplexing	FASTX-Toolkit	http://hannonlab.cshl.edu/fastx_toolkit
	Stacks	(Catchen et al., 2011)
	Trimmomatic	(Bolger et al., 2014)
	TrimGalore	http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
General Aligners, Psuedoaligners, De novo Annotators	Bowtie2	(Langmead and Salzberg, 2012)
	Kallisto	(Bray et al., 2016)
	Novalign	http://www.novocraft.com
	SOAP2	(Li et al., 2009b)
	STAR	(Dobin et al., 2013)
	Tophat2	(Kim et al., 2013a)
Post-Alignment Processing, QC, Counting, and Visualization	htseq-count	(Anders et al., 2013)
	IGV	(Thorvaldsdottir et al., 2013)
	RNA-SeQC	(DeLuca et al., 2012)
	RSeQC	(Wang et al., 2012a)
	Rsubread	(Liao et al., 2013)
	SAMtools	(Li et al., 2009a)
Differential Expression Analysis	edgeR	(Robinson et al., 2010)
	DESeq2	(Anders et al., 2013)
	baySeq	(Hardcastle and Kelly, 2010)
	Cuffdiff2	(Trapnell et al., 2013)
	DEGseq	(Wang et al., 2010a)
	EBSeq	(Leng et al., 2013; Leng et al., 2015)
	voom	(Law et al., 2014)

Table 1.2: RNAseq preprocessing and analysis tools currently in common use

during sequencing. Trimming, filtering, and demultiplexing can be accomplished with the command line or a tool like the FASTX-Toolkit (http://hannonlab.cshl.edu/fastx_toolkit) or Stacks (Catchen et al., 2011). For a more thorough review of QC metrics see (Li et al., 2015)).

(ii) *Align reads to reference genome*: Aligners generally fall into two basic categories - those that emphasize speed or those that emphasize sensitivity. Fast aligners include STAR (Dobin et al., 2013), Bowtie2 (Langmead and Salzberg, 2012), and SOAP2 (Li et al., 2009b). More sensitive aligners include Novalign (http://www.novocraft.com) and SHRiMP2 (David et al., 2011). TopHat2 (Kim et al., 2013a), is one of the most widely used alignment tools though it has been replaced by HISAT (Kim et al., 2015) which is more accurate and faster. A step-by-step guide to Tophat2 usage can be found in (Anders et al., 2013). For a comparative list of aligners see (Fonseca et al., 2012; Engstrom et al., 2013; Baruzzo et al., 2017). The output of aligned reads is usually stored in either SAM or BAM file formats which can be processed with SAMtools (Li et al., 2009a). The SAM format is a readable formatted file that can be examined visually. The BAM format is a compressed SAM file that can be processed much more quickly.

(iii) *Generate BAM statistics, visualize aligned reads*: At this stage, it is important to determine the percentage of mapped reads (generally greater than 75% for a successful alignment). Duplicate reads may be identified but the general consensus is that filtering out these reads can bias downstream DGE analysis. The percent of rRNA reads present is also an important metric. Theoretically, the rRNA removal step in library preparation should have removed it but this step is not 100%

efficient. All of the above statistics can be generated either using SAMtools, RSeQC, or RNA-SeQC. The next step is visualization of the alignment. Although SAM files are readable, it is much easier to use a genome visualization tool to confirm the success and quality of the alignment. There are many tools available for this purpose such as IGV (Thorvaldsdottir et al., 2013). SAMtools can be used to convert SAM files into position/index-sorted BAM files so they can be visualized. Visualization should be used to confirm the success of the alignment such as correct mapping over the exon-exon junctions. As a final QC check, if certain transcripts are already known to have differential expression, the genome visualizer may be used to confirm that these expression differences are seen in the aligned RNA samples.

(iv) Obtain raw read counts and normalize: Now that the RNA has been aligned to the genome, raw read counts can be generated for each transcript. HTSeq can be used for this purpose using the htseq-count tool (Anders et al., 2015). In the R/Bioconductor environment, the Genomic Features and Genomic Alignment packages can be used (Lawrence et al., 2013). Typically, the next stages of analysis will be the comparison of differentially expressed genes. The issue is that raw counts cannot be compared to each other, either within the same library or comparing samples between different libraries, without subsequent normalization. This is because of an inherent bias in the sequencing process that results from either the depth of sequencing or the length of a transcript. If two transcripts actually have equal expression but the second transcript was sequenced to a greater depth, the raw counts would indicate that the second transcript has a greater differential expression when in reality they are the same. If two transcripts have equal

expression but the second one is longer than the other, higher raw counts would also be reported when in reality the counts of both transcripts should be the same. To overcome this issue, a number of different normalization strategies have been developed (Dillies et al., 2013; Zyprych-Walczak et al., 2015; Lin et al., 2016). One of the first strategies employed to overcome this issue a normalization called RPKM - reads per kilobase transcript per million reads (Mortazavi et al., 2008). FPKM was introduced later for paired-end data and employs the same principle except that it accounts for the fact that two reads will be mapped to a fragment. RPKM is still the most widely used method despite the fact that it does not take into account transcript length resulting in a bias in the variance of gene expression estimates (Oshlack and Wakefield, 2009). RPKM and FPKM were later expanded to other methods that divide raw counts by either the median, quantiles (Law et al., 2014), total counts, or upper quantile (see review cited above for a comparison). Two alternative methods employed by the R packages DESeq2 (Love et al., 2014) and edgeR (Robinson et al., 2010) take different approaches entirely based on the assumption that most genes are not differentially expressed.

Downstream data analysis

In order to compare normalized counts between different groups of genes the distribution of the data must be determined. Since counts from RNAseq data are discrete rather than continuous (even after a transformation) only specific distributions can be used to model the data. The negative binomial model is currently popular because it corrects for errors that result from modeling with the Poisson distribution and is used by both the DESeq2 and EdgeR packages. There are a number of other software options available for

DGE analysis that use different types of modeling and no clear consensus on what approach is the best (Rapaport et al., 2013; Soneson and Delorenzi, 2013; Seyednasrollah et al., 2015). Sonenson and Delornzi have provided a detailed analysis of the different methods and their performance depending on the features of the data (sample size, degree of differential expression etc.). It would be prudent to choose the method of choice based on this information and even try a few different methods to see how different the generated differential expression values are for each gene.

In addition to DGE analysis, there are a number of options available to further analyse the data. Here we briefly list a few approaches. After obtaining the list of protein and RNA, the data may be subjected to gene ontology (GO) clustering to find broad associations in the data. There are a number of GO clustering software available but DAVID is among the simplest to use (Huang da et al., 2009a; 2009b). Given a specific list of genes, DAVID will associate the data with specific gene ontologies in the hierarchy and cluster the most representative terms. The full GO ontology as well as smaller version are available. DAVID also includes annotation for pathways, protein domains, and protein interactions providing a convenient means get a broad representation of the data. GOseq is another alternative which takes into account the transcript length bias mentioned above (Young et al., 2010).

Network analysis can be extremely helpful in visualizing complex data. Cytoscape and its associated plug-ins give the user a high amount of customization and flexibility to visual map connections between genes associated in different pathways, molecular interaction networks, and disease databases (Shannon et al., 2003). Some plugins also allow input of gene expression data as an extra visual dimension to the data. For example, the Enrichment Map plugin from the Bader Lab allows the user to take GO clustered data from DAVID and visualize it as a network. Enrichment Map helps to simplify data

produced by DAVID in a visual format for easier interpretation (Merico et al., 2010; Merico et al., 2011). This network-based approach is especially useful for modeling interaction networks of proteomics data and has been used to characterize molecular networks in synaptic plasticity (Pocklington et al., 2006).

Clustering (K-means, hierarchical clustering etc.) is also used to classify gene expression data. Clustering will group a set of genes into categories based on the similarity of their expression levels. Used in tandem with a heatmap, expression changes can be visualized on a color scale while simultaneously grouped by similarity in expression level changes (D'Haeseleer, 2005). The Genesis software[2] provides an easy way to perform various clustering methods. Many clustering strategies assume the data is normally distributed which is not usually the case for RNAseq data. One option is to apply a transformation to the data so that the counts are more normally distributed and then apply the method (Zwiener et al., 2014). Alternatively, there are a few methods that have been recently developed that apply clustering strategies to either Poisson or negative binomial distributions which much more closely approximate RNAseq data (Witten, 2011; Si et al., 2014). Principal components analysis (PCA), support vector machines (SVM) and other tools for dimensionality reduction and classification analysis can also be applied to RNAseq data (Tan et al., 2014).

COMBINED TRANSCRIPTOMICS AND PROTEOMICS

The overall goal of RNA sequence identification and bioinformatics analysis on synaptic fractions is to determine how the synapse is remodeled under dynamic physiological and disease conditions. Since expression of proteins with coordinated

[2] http://genome.tugraz.at/genesisclient/genesisclient_faq.shtml

functions can rapidly change the efficacy of a synapse, proteomics analysis is thus a critical counterpart to RNA sequencing experiments (Martin and Zukin, 2006; Zukin et al., 2009; Fernandez-Moya et al., 2014; Rosenberg et al., 2014). The increased sensitivity and accuracy of high-throughput methods have now made it feasible to conduct RNASeq studies and follow up by mass spectrometry. While the application of Omics to synaptic data is slowly growing (Collins et al., 2006; Fernandez et al., 2009; Geschwind and Konopka, 2009; Fritzsche et al., 2013; Ch'ng et al., 2015; Hussain and Bashir, 2015; Broek et al., 2016; Kenney et al., 2016; Loos et al., 2016; Niere et al., 2016) few studies combining multiple Omics have been performed in a synaptic context (Valor and Grant, 2007; Kitchen et al., 2014). To date, combined approaches have been used to determine differences in neuronal cell types (Sharma et al., 2015), the effects of oxidative stress in synaptosomes (Flynn et al., 2012), the age-specific differences in brains of old and young rats (Ori et al., 2015), differences in stages of embryonic development (Hartl et al., 2008), cell type-specific proteins enriched in the brain over other tissues (Sjostedt et al., 2015), and the molecular dynamics of Rett Syndrome (Li et al., 2016). One of the most important findings from the combined omics approach is the confirmation that protein and RNA levels from the same tissue or single-cell do not always correlate (Maier et al., 2009; Olivares-Hernandez et al., 2011; Haider and Pal, 2013). Therefore, such an integrated approach will allow more accurate modeling of the dynamic molecular interplay that underlies synaptic function such as post-transcriptional regulation by non-coding RNA as well as proteomic regulation through post-translational modifications. In the next few sections, we review a number of possible assays to be used for isolating protein populations as well as *de novo* protein synthesis assays.

Isolation of proteins populations

Proteins play a vital role in synaptic plasticity as they are part of the molecular substructure that underlies neuronal circuitry. Many events in local translation are mediated by proteins, particularly kinases, that are able to fine-tune the timing and specific expression of mRNAs through the regulation of ribosome and other localized events in the dendrite (Giese and Mizuno, 2013; Katche et al., 2013; Rosenberg et al., 2014; Santini et al., 2014). As many of these proteins interact with each other in specific ways to affect synaptic efficacy by altering the molecular composition of the dendritic spine, particularly in response to cellular signaling events at the postsynaptic membrane or nuclear signals, it is vital to understand these interactions in greater detail. Proteomic characterization of synaptic plasticity is ongoing and some recent studies have used system biology approaches to analyze the resulting data (Collins et al., 2006; Fernandez et al., 2009; Niere et al., 2016). In this section, we will consider a number of techniques that may be used to isolate protein populations before high-throughput analysis via mass spectrometry (Table 1.3).

Co-Immunoprecipitation (Co-IP) and pull-downs

The Co-IP is a method for determining protein-protein interactions; while an immunoprecipitation (IP) is intended to isolate a single molecule of interest, a Co-IP is able to isolate the entire complex (”prey”) the protein of interest (”bait”) is bound to (Free et al., 2009). Thus, specific modifications must be made to the general IP procedure to ensure the bait and prey interaction is not lost during the procedure. Similar to the RIP procedure, the lysates are combined with an antibody specific to a protein of interest and a scaffold (usually a magnetic or agarose bead) which has affinity to the antibody. Collecting the bead-antibody-protein complex allows one to selectively isolate the protein from the lysate

attached to any other proteins that it was bound to *in vivo*. The protein bound to the protein of interest can then be collected. This procedure is easier to perform than RIPs as proteins are less prone to degradation than RNA. However, it is still limited by antibody availability and non-specific binding issues inherent to antibody-based methodology and other technical limitations (Markham et al., 2007). Pull-downs have the same goal as the Co-IP but do not use an antibody for the isolation of the bait and prey (Free et al., 2009). Instead, the bait protein is isolated either through a bioengineered molecular tag or through some kind of covalent attachment to the bead. Though this method does not require an antibody it is still limited by the ability to engineer and express the bait containing the tag in molecularly accessible position. Following isolation by either method, the samples can then be prepared for mass spectrometry analysis to identify the prey proteins. Recently, an alternative beads-free method has been developed in which protein are isolated in wells of a microplate coated in either protein A or G (Mikula et al., 2015). The authors note that this method allows multiple IP reactions to be performed simultaneously allowing better control of experiment variability, can be less costly than traditional IP methods, can be automated using robotics, and can be adapted to thermocyclers for protein denaturation steps. The drawback to these methods is that they fail to catch weak or transient protein-protein interactions due to the stringent bead washing steps. This issue may be mitigated by formaldehyde crosslinking (see below).

Labeling methods

Label transfer protein interaction analysis utilizes a crosslinking reagent to join the bait and prey proteins. A new variation of this technique which does not rely on covalent

Assay Type	Technique	Advantages	Disadvantages	Reference
Direct RNA-protein interactions	Immunoprecipitation	Single molecule (IP) or complex (Co-IP)	Antibody-based, non-specific binding, difficult to detect proteins with low expression, cannot identify transient interactions	(Markham et al., 2007; Sutherland et al., 2008; Free et al., 2009; Mikula et al., 2015)
	Pull-down	Tag-based, does not require antibody	Tag may be difficult to engineer and may alter protein function, difficult to detect proteins with low expression, cannot identify transient interactions	
Labelling methods	Label transfer protein interaction	Isolation of transient protein-protein interactions, interaction within physiological context	Difficult to balance dissociation timing with label transfer molecule; alternative approach in Liu et al. 2009	(Liu et al., 2007)
	BioID	Overcomes difficulties in Co-IP and pull-downs	Best suited for culture work, fusion protein may interfere with protein interactions, biotin may alter properties of protein/interacting partners	(Roux et al., 2012)
Crosslinking methods	Analysis of oligo(dT)-purified mRNPs	Can observe dynamic changes in RNA-protein interactions	Cannot identify non-Poly(A) proteins or microRNAs	(Baltz et al., 2012)
In vitro binding	Modified phage display	Compliment to other protein-protein interactions methods	Technically challenging, will pick up non-physiological interactions	(Di Niro et al., 2010)
	RNA bait quantitative proteomics	Compliment to other protein-protein interactions methods,	Will pick up non-physiological interactions	(Butter et al., 2009)
	Size-exclusion quantitative proteomics	Identify transient interactions, does not require any tags	Will pick up non-physiological interactions	(Kirkwood et al., 2013)

Table 1.3: Summary of methods for identifying protein-protein interactions

modification of the bait protein has also been developed (Liu et al., 2007). The crosslinker is labeled, for example, with biotin or a fluorescent marker. Next, the crosslink is cleaved leaving the labeled end attached to the prey proteins. The biotin-labeled prey proteins can then be isolated and analyzed by mass spectrometry. A major advantage of this technique is its ability to isolate transient protein-protein interactions. Another approach called BioID utilizes a biotin-ligase-fused bait protein to isolate interacting prey proteins *in vivo* after adding biotin to cell culture media (Roux et al., 2012). The biotin-ligase is then able to add biotin to lysine residues on neighboring proteins which can then be isolated and analyzed by mass spectrometry. This technique is best suited for cell culture work rather than tissue.

Crosslinking methods

Crosslinking approaches are now common in high-throughput RNP isolation techniques. However, this approach is also adaptable to the study of protein-protein interactions. Crosslinking methods can also be applied to Co-IPs though this technique has not yet been fully optimized or tested (Sutherland et al., 2008). The analysis of oligo(dT)-purified mRNA ribonucleoproteins (mRNPs) method combines a protein and RNA approach to allow for the identification of RNPs as well as the mRNAs they are bound to. First, RNA is labeled with photoreactive nucleoside analogs of uridine and guanosine, allowing selective identification of RNA over DNA and increased crosslinking efficiency. Next, the protein-RNA complexes are crosslinked and oligo(dT)s are affinity purified and the protein and RNA can then be separately identified by mass spectrometry and RNA sequencing respectively (Baltz et al., 2012). Though application of this technique will be useful for the study of mRNPs, it will miss nonpolyadeylated RNAs such as microRNAs

whose importance in learning and memory is very recently being understood (Wang et al., 2012b).

In vitro binding methods to determine protein-protein or RNA-protein interactions

An alternative strategy for discovering protein-protein interactions was developed by Sblaterro and colleagues that utilizes phage display, an *in vitro* method similar to yeast two hybrid (Di Niro et al., 2010). First, polyadenylated RNA is fragmented and reverse transcribed into cDNA. Next, the cDNA is cloned into bacteria and an ampicillin-based selection method is used to remove nonfunctional clones that do not encode ORFs. The vectors are ligated into a bacteriophage minor coat protein gene allowing the phage to "display" the protein on its outer surface. Now, a purified protein of interest can be introduced to see if they can bind to any of the proteins displayed by the phages. The cDNA from the phages recovered from this process is then sequenced to determine which phage-displayed proteins interacted with the protein of interest. Additionally, an interesting technique has been developed that allows for the identification of putative RBPs (Butter et al., 2009). This technique works in reverse of a RIP, instead using the RNA as the bait to determine potential RBP binding partners. An RNA bait sequence of interest is coupled to an S1 aptamer. These labeled RNA fragments are then introduced to a cell population. RNA-protein complexes are then pulled-down and analyzed by quantitative mass spectrometry. This procedure is a powerful technique to complement RNA-based procedures. Lastly, a size-exclusion chromatography approach combined with quantitative proteomics useful for predicting protein-protein interactions (Kirkwood et al., 2013). First lysates are sonicated in PBS followed by size-exclusion chromatography which separates on the basis of both size *and* shape while maintaining cellular protein-protein interactions.

This results in 40 different fractions where the larger fractions represent larger interacting protein complexes down to the smallest fractions which have less interacting partners. The fractions are then evaluated by mass spectrometry to determine interacting partners of every protein in the lysate. Each protein will have a specific elution profile across all the fractions to demonstrate fraction behavior for each protein. Proteins showing similar coincidence in their elution profile for different fractions are thus considered to be associated. Like many *in vitro*-based methods, these techniques may capture some interactions that do not occur *in vivo*. Nevertheless, it may serve as a useful tool for filtering false positives in other techniques and determining possible interaction partners for a protein for further study. The size-exclusion and quantitative analysis approach may be especially useful as a validation tool for Co-IPs which have a tendency to pick up non-physiological interactions. Such approaches may be able to filter out these cases by demonstrating that such interactions are unlikely to occur in biological systems.

De novo protein synthesis assays

The separation of established protein from de novo protein synthesis has been challenging. Methods such as BONCAT (bioorthogonal non-canonical amino acid tagging) were recently developed as a means of selectively isolating newly synthesized proteins from a larger population (Dieterich et al., 2006; Dieterich et al., 2007). Utilizing click chemistry, newly-synthesized proteins incorporate a non-canonical amino acid using the cell's own translational machinery which can be conjugated to biotin. One of the great advantages of this method is that endogenous proteins are labeled as opposed to introducing exogenous reporter proteins. These proteins can then be isolated from the rest of the population using affinity chromatography and analyzed by mass spectrometry. Instead of

non-canonical amino acids, SILAC (stable isotope labeling by amino acids in cell culture) uses heavy or light isotopes of arginine and lysine (Ong et al., 2002). One population of cell culture will be grown with media containing the heavy isotope and a second population will use the light isotope. The proteins from both populations are then combined and analyzed by mass spectrometry. A protein of interest can then be compared between the two different cell cultures – one will be heavier due to incorporation of the heavy isotope and its abundance levels can be compared to the same protein in the other cell culture thus demonstrating a change in expression between two cellular conditions.

Though both BONCAT and SILAC address different kinds of questions, there have been a number of recent developments of SILAC such as pulsed SILAC (pSILAC) that allows for the determination of newly synthesized proteins (Chen et al., 2015). There have also been a few recent attempts to combine BONCAT and SILAC together (Genheden et al., 2015; Kenney et al., 2016). Kenny et al. applied this technique to cell culture by first introducing medium or heavy arginine to the cells (SILAC) followed by the addition of a non-canonical amino acid to isolate newly synthesized proteins (BONCAT)(Kenney et al., 2016). This combined approach allowed for the comparison of specific newly synthesized proteins rather than the entire population. Furthermore, another approach called BONLAC combines the two methods and optimizes BONCAT conditions so that it may be performed in intact brain slices (Bowling et al., 2016). In addition to labeling with a non-canonical amino acid (AHA, a methionine analog) as in BONCAT, Bowling et al., also labelled with medium or heavy arginine thus allowing them to selectively isolate de novo synthesized proteins within a short time window. An in vivo BONCAT approach was also developed by Liu and Cline which allowed them to assess the effect of an FMRP knockdown on protein synthesis-dependent behavioral plasticity in the Xenopus visual system (Liu and

Cline, 2016). The evolution of this method has greatly expanded our knowledge as it now allows the detection of activity-dependent de novo protein synthesis *in vivo*.

FROM STATISTICAL MODELS TO SCIENTIFIC MODELS: EXPERIMENTAL VALIDATION AND VISUALIZATION OF IDENTIFIED PROTEINS AND RNA

One of the greatest advantages of performing unbiased screens such as RNAseq and mass spectrometry is the identification of novel interactions between mRNAs and/or proteins. The wealth of data generated from such high-throughput experiments can be used to build general models to guide the direction of future scientific research. For example, constructing protein-protein interaction (PPI) networks from proteomics data can be used to guide experimental research leading to new target identification in animal models of disease (for further explanation of PPI network construction and analysis see (Raman, 2010)). For example, this technique was recently used by Niere et al., with the reasoning that many individuals who have neurological disorders with dysregulated protein synthesis due to overactive mTOR signaling also suffer from epilepsy (Crino, 2008; Pun et al., 2012; Brewster et al., 2013; Wong, 2014; Sosanya et al., 2015b; Niere et al., 2016). To identify common proteins associated with epilepsy, Alzheimer's disease, and Autism Spectrum Disorders, all disorders with overactive mTOR signaling, a PPI network was established that identified 5 "hub proteins" based on its high level of connectivity with other proteins in the network. One hub protein was Parkinson Protein 7 (Park7 or DJ-1), a protein that has many functions but most recently has been identified as an RBP (van der Brug et al., 2008). Importantly, the investigators went on to show that Park7 protein synthesis was regulated by mTOR and that it is overexpressed at synapses in a mouse model of Tuberous Sclerosis Complex, a form of ASD with overactive mTOR. Below we review techniques that may be used to visualize the localization and relative quantity of protein and/or RNA

to validate models generated from high-throughput methodologies (Table 1.4, Figure 1.5). While most of these techniques suffer from the single protein/RNA approach, they do provide a powerful means of validating findings that may lead to new target identification of diseases with dysregulated protein synthesis.

Visualization and quantification of RNA localization

In situ hybridization

The first requirement for a protein to be synthesized in dendrites is that the mRNA coding for that protein is either targeted to the dendrites in response to activity or constitutively resides in dendrites. Some of the earliest work in this field utilized various methods such as in situ hybridization (ISH) and microarrays in conjunction with subcellular fractionation to catalog the RNAs present at the dendrite. One of the greatest advantages ISH is that it does not rely on subcellular fractionation – the transcript can be identified in fully intact tissue or cell culture. Thus, ISH and its related methods are very useful for confirming results obtained from subcellular fraction preparations. ISH has been used to characterize the localization and distribution of the RNA transcripts for the calcium/calmodulin-dependent protein kinase II alpha (CaMKII), microtubule-associated protein 2 (MAP2) and fragile X mental retardation protein (FMRP), some of the earliest localized transcripts discovered in the field (Hinds et al., 1993; Paradies and Steward, 1997; Miyashiro et al., 2003). The development of fluorescence in situ hybridization (FISH) has allowed researchers to visualize distribution patterns and quantify the amount of localization at a much higher resolution. There have been significant technical improvements made to FISH over the years which provide greater resolution and better quantification (Swanger et al., 2011). The best example of how technical improvements in

Detection of a reporter or endogenous RNA/protein	Technique	Detection Chemistry	Detection Type	Selected References
Endogenous (RNA)	(F)ISH	Hybridization probe; fluorescent antibody	RNA localization	(Lecuyer et al., 2007; Swanger et al., 2011; Cajigas et al., 2012)
	Single RNA tracking	Fluorescent dyes	RNA localization and translation	(Katz et al., 2016)
	qRT-PCR	Fluorescent DNA intercolator	Relative RNA quantitation	(Raymaekers et al., 2009)
	NanoString	Hybridization probe	Relative RNA quantitation	(Kulkarni, 2011; Cajigas et al., 2012)
Endogenous (Protein)	BONCAT	Biotin	De novo synthesis	(Dieterich et al., 2006; Dieterich et al., 2007)
	SILAC	Heavy/light chain amino acid isotopes	De novo synthesis	(Ong et al., 2002)
	BONLAC	Biotin + heavy/light chain amino acid isotopes	De novo synthesis	(Bowling et al., 2016)
	FUNCAT	Biotin, fluorescent antibodies	De novo synthesis	(Dieterich et al., 2010; Tom Dieck et al., 2012; Kos et al., 2016; Liu and Cline, 2016)
	BONCAT/FUNCAT-PLA	Biotin, fluorescent antibodies	De novo synthesis	(tom Dieck et al., 2015; Workman et al., 2015; Niere et al., 2016)
	SUnSET	Fluorescent antibody	De novo synthesis	(Schmidt et al., 2009; Batista et al., 2016)
Reporter construct (Protein)	Destabilized GFP (dGFP)	Fluorescent protein (fusion construct)	De novo localized synthesis	(Li et al., 1998; Aakalu et al., 2001)
	Kaede	Photoactivatible GFP-like fluorescent protein (PAFPs)	De novo localized synthesis	(Ando et al., 2002; Lukyanov et al., 2005; Raab-Graham et al., 2006; Workman et al., 2015)
	Dendra2	Photoactivatible GFP-like fluorescent protein (PAFPs)	De novo localized synthesis	(Gurskaya et al., 2006; Chudakov et al., 2007; Wang et al., 2009; Lee et al., 2011)
	Venus fluorescent reporter	Fluorescent protein (fusion construct)	De novo localized synthesis	(Tatavarty et al., 2012; Ifrim et al., 2015)
	Biarsenical probes (FlAsH, ReAsH)	Fluorescent dyes	De novo localized synthesis	(Ju et al., 2004; Martin et al., 2005; Rodriguez et al., 2006)
	TimeSTAMP	Epitope tag	De novo localized synthesis	(Lin et al., 2008; Lin and Tsien, 2010)
	MiniSOG	Fluorescent protein (fusion construct)	De novo localized synthesis	(Shu et al., 2011)
	Luciferase Flash Kinetics	Luciferase (fusion construct)	De novo localized synthesis	(Na et al., 2016)
	SINAPS	Fluorescent protein (fusion construct)	Real-time translation dynamics	(Wu et al., 2016)

Table 1.4: Visualization and detection techniques for RNA and protein downsream of high-throughput experiments

FISH have changed dogma was first described by Krause and colleagues using mRNA localization in the Drosophila embryo as a model. Prior to this study the estimated number of mRNAs that had distinct subcellular localizations in the embryo was ~1 -10%. Lecuyer et al., screened roughly 25% of the genome and found ~ 71% of mRNAs screened had unique subcellular distribution patterns. The authors went on to suggest that since most of the transcribed mRNAs have a distinct subcellular localization these data imply that most cellular processes are mediated through mRNA localization (Lecuyer et al., 2007). Importantly, a similar finding was suggested by Schuman and colleagues in neuronal dendrites. They also employed FISH to verify high-throughput RNA sequencing data (Cajigas et al., 2012). Herein they identified 8,379 transcripts. By subtracting transcripts related to glial cells, interneurons, nuclei, blood vessels, and mitochondria, they suggest that 2,550 transcripts are localized in axons or dendrites. High-sensitivity FISH was then used to validate the localization of 50 of these transcripts. This work identified many known synaptic transcripts as well as newly discovered transcripts that had not been previously detected. In light of these findings synapse remodeling of synapses during synaptic plasticity is due to dendritic over somatic mRNA translation (Cajigas et al., 2012). Finally, the a more recently modification to FISH allows for the visualization of single RNA molecules at high resolution. Singer and colleagues have developed a method that allows the tracking of single RNA molecules during translation. Thus, the researchers labeled ribosomes and mRNA molecules and correlated their signals to determine if the mRNA was undergoing active translation (Katz et al., 2016). Thus, collectively RNA sequencing data combined with high resolution FISH, for the first time, is allowing investigators to catalog mRNAs localized to site specific dendritic compartments leading to new testable hypothesis regarding memory allocation for information storage.

Quantification of mRNA

While FISH by itself provides subcellular localization of specific mRNAs, quantification of mRNAs in the soma versus dendritic compartments has been challenging. The development of the NanoString nCounter Gene Expression Assay has helped resolve that issue (Kulkarni, 2011). NanoString enables the detection and quantification of up to 800 mRNA molecules (without any conversion to cDNA or PCR amplification) using colored probe pairs. Cajigas et al. used NanoString technology to answer the long-standing question of whether select mRNAs segregate, with some enriched in the soma and some enriched in dendrites and axons (Cajigas et al., 2012). While this finding had been suggested by quantitative RT-PCR comparing hippocampal synaptosomal mRNA to total lysate mRNA (Raab-Graham et al., 2006) the concern of somatic contamination in the synaptosomal fraction tempered the interpretation of this finding. Still, in cases when mRNA levels are low-abundance or quick validation is a necessity, qRT-PCR is remains a valid method for confirming RNAseq or microarray data, or quantifying relative abundance levels of a specific transcript between samples (Raymaekers et al., 2009).

Visualization of protein localization and new translation

SUnSET, FUNCAT, and FUNCAT/BONCAT-PLA

Labeling newly synthesized proteins has been used to validate high-throughput studies investigating local translation (Buxbaum et al., 2015a; Buxbaum et al., 2015b). Since many of these techniques rely on fluorescent microscopy, it is possible to visualize and distinguish new and old proteins within a particular neuronal compartment. Surface sensing of translation (SUnSET) utilizes puromycin, a ribosome elongation inhibitor and aminoacyl-tRNA synthetase analog, to monitor the translation of de novo protein synthesis

(Schmidt et al., 2009). SUnSET uses monoclonal antibodies to detect the incorporation of puromycin into the polypeptide chain during translation. Subsequent detection and visualization by either flow cytometry or fluorescent microscopy thus indicates new protein synthesis.

Single-molecule imaging of nascent peptides (SINAPS) is another alternative approach for monitoring localized transcripts in vivo which allows one to observe the real-time translation of nascent mRNA molecules at different translational stages (Wu et al., 2016). In order to strengthen the relatively weak signal of the nascent peptide and to distinguish it from background, SINAPS draws upon SunTag, another recent method that allows for the amplification of fluorescent intensity by recruiting multiple copies of GFP to a target protein (Tanenbaum et al., 2014).

The FUNCAT (fluorescent non-canonical amino acid tagging) assay is similar to the BONCAT technique described earlier which labels newly synthesized proteins using noncanonical amino acids and click chemistry to conjugate biotin. FUNCAT, however, uses fluorescently-tagged amino acids to allow for the identification of newly synthesized proteins in situ by fluorescence microscopy (Dieterich et al., 2010; Tom Dieck et al., 2012; Kos et al., 2016). The drawback to FUNCAT is the fact that while it can identify whole populations of de novo synthesized proteins, it cannot identify specific ones. To overcome this difficulty, a modification was developed using a proximity ligation assay (PLA). The use of FUNCAT in tandem with PLA allows for the identification of a newly synthesized protein of interest (tom Dieck et al., 2015). FUNCAT-PLA – also called BONCAT-PLA in (Niere et al., 2016) – utilizes a biotin antibody to identify a protein of interest that has incorporated a non-canonical amino acid, signifying new protein synthesis. Another antibody is used to identify the protein of interest itself. When the secondary antibodies (pre-conjugated to a specific oligonucleotide sequence) for the two primary antibodies are

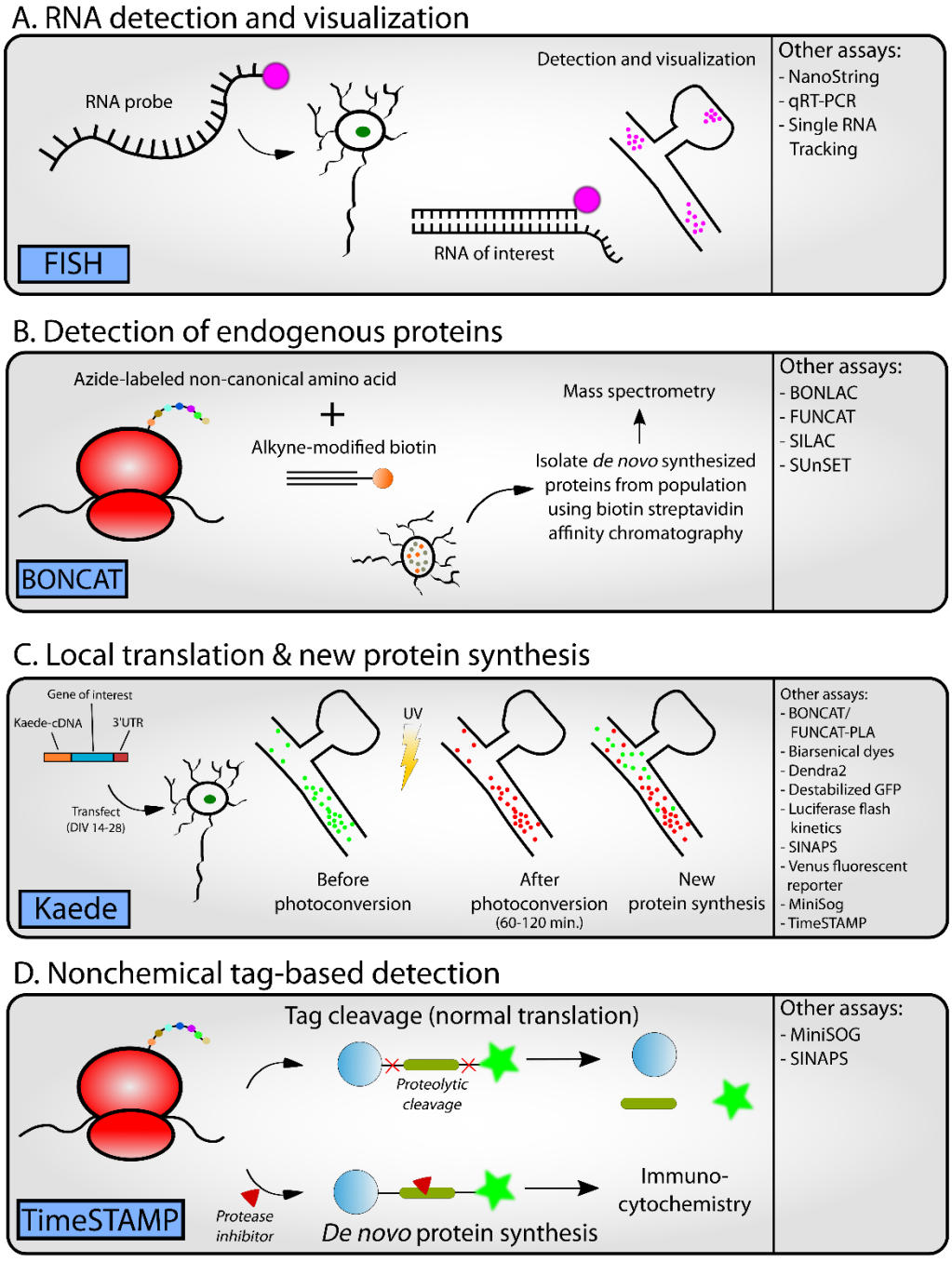


Figure 1.5: Workflow for experimental follow-up

Following the completion of high-throughput methodologies, researchers can perform a variety of different types of experiments to follow-up predictions seen in the data. Representative assays are depicted in cartoon form (A.-D., *left panel*) and alternative approaches (A.-D., *right panel*) or

in close enough proximity, they are ligated together and undergo rolling circle replication in the presence of fluorescent nucleotides to produce a signal. Alternatively, the Puro-PLA assay utilizes puromycin, a molecule that disrupts translation resulting in the release of the newly-synthesized protein which can then be identified with an anti-puromycin antibody (Buhr et al., 2015; tom Dieck et al., 2015).

Recently, FUNCAT/BONCAT-PLA has been used to assess the role of mTOR complex 1 (mTORC1) in synapse modification in vivo. mTORC1 is a protein complex that regulates local dendritic translation (Tang et al., 2002; Cammalleri et al., 2003; Stoica et al., 2011). In a recent study, Niere et al. performed mass spectrometry on different subcellular fractions of neurons derived from rat cortices after an intraperitoneal injection with rapamycin, an inhibitor of mTORC1. Notably, the mass spectrometry reports changes in protein expression and does not differentiate between protein synthesis and protein

---

references can be found in Table 1.4 or the text. **A.** RNA detection and visualization using fluorescence in situ hybridization (FISH) in neurons. An RNA probe complementary to an RNA of interest is designed. A fluorescent tag is added to the probe for detection downstream. The probe is introduced to a neuron population. Once the RNA enters the neuron, it binds to the RNA of interest. Cells can then be fixed and visualized by fluorescence microscopy to visualize the location of the RNA within a tissue or cell culture system. **B.** Detection of endogenous proteins using bioorthogonal non-canonical amino acid tagging (BONCAT). An azide-labeled non-canonical amino acid is introduced to a cell population. During translation, this amino acid is conjugated to alkyne-modified biotin using click chemistry, thus tagging a newly-synthesized protein. Affinity chromatography is used to isolate these biotin-labeled de novo synthesized proteins from the greater protein population. The isolated proteins can be identified using mass spectrometry. **C.** The photoconvertible fluorescent protein Kaede can be used to visualize local translation/new protein synthesis of a protein of interest. A vector containing Kaede-cDNA is introduced to a cell culture population. Prior to photoconversion, Kaede fluoresces green revealing the current population of a protein of interest. After application of UV light, Kaede fluoresces red. Any further green signal that appears later is indicative of new translation of the protein of interest. **D.** Options for non-chemical tag-based detection include time-specific tagging for the age measurement of proteins (TimeSTAMP). A fusion construct is generated between the protein of interest and an epitope tag flanked by a cassette for the hepatitis C virus protease. After translation, the protease cleaves itself and the epitope tag away from the protein of interest. The protease inhibitor BILN-2061 may be added to the cells at any time to inhibit the proteolytic cleavage. Thus, the protein of interest can retain its tag. Using this system, researchers can separate new and old protein synthesis after a specific time point.

stability. To differentiate between these two cellular mechanisms, they used FUNCAT/BONCAT-PLA to determine if new protein synthesis of select candidates identified by mass spectrometry (Niere et al., 2016). Indeed, FUNCAT/BONCAT-PLA demonstrated that new protein synthesis levels of Snap25 and Gap43 altered in response to treatment with the mTOR inhibitor rapamycin, and were consistent with the observed changes in protein expression indicated by mass spectrometry. Thus, new protein assays such as FUNCAT/BONCAT-PLA can detect new protein synthesis, providing mechanistic detail that mass spectrometry hints at but does not confirm.

Destabilized GFP (dGFP)

While FUNCAT/BONCAT-PLA is useful to detect new protein synthesis it does not provide direct evidence for new protein synthesis in dendrites. For this reason, fluorescent translation reporters fused to dendritic targeting sequences of the proteins of interest still remains the best way to visualize new protein synthesis in dendrites. This approach was first developed by Schuman and colleagues (Aakalu et al., 2001). Destabilized-GFP (dGFP) was developed to address experiments that required a fluorescent reporter for proteins with transient expression (Li et al., 1998). Schuman and colleagues capitalized on the rapid turnover properties of dGFP and created a reporter consisting of cDNA coding for a myristoylated dGFP fused between the 5' and 3'UTRs of CaMKIIα mRNA. They reasoned that the inclusion of the UTR sequences in their reporter construct would ensure that the mRNA targets the dendrite, the addition of a myristoylation sequence (myr) tethers the reporter to the membrane and thus prevents diffusion, and after photobleaching the neuron, new GFP signal detected in the dendrite is due to mRNA translation. To ensure this was the case, they continuously photobleached the soma so that any new protein synthesized in the soma would not be detected in the dendrites. One caveat

that may hinder this assay is if the protein dGFP is fused to has a higher stability than dGFP, thus preventing the rapid turnover of GFP. In spite of this limitation, myrdGFP fused to the appropriate targeting sequences can be used to investigate compartment-specific translation (Aakalu et al., 2001).

Kaede, Dendra2, Venus, and Biarsenical probes

An alternative approach is the usage of photoactivatable GFP-like fluorescent proteins (PAFPs) (Lukyanov et al., 2005) which overcomes the limitation of variability in mRNA stability. Kaede is a protein that allows the separation of new and old translation on the basis of UV-induced photoconversion (Ando et al., 2002). Kaede can be fused to specific proteins with the appropriate dendritic targeting sequences and used to report new translation by comparing mean puncta intensity before and after photoconversion (Leung et al., 2006; Raab-Graham et al., 2006; Leung and Holt, 2008; Banerjee et al., 2009; Workman et al., 2015). Kaede's tetrameric structure has a distinct advantage when it comes to the tagging of ion channels which are typically tetramers (Raab-Graham et al., 2006). For example, Kaede has been used to monitor the local translation of Kv1.1, an ion channel whose translation is regulated by a microRNA, demonstrating its use and application as a detector of new protein synthesis (Sosanya et al., 2013). The Dendra2 reporter is another photoconvertible fusion protein that functions similarly to Kaede (Gurskaya et al., 2006; Chudakov et al., 2007; Wang et al., 2009; Lee et al., 2011). Dendra2 has a monomeric structure which increases functional capabilities since it can be used in protein fusion constructs. For example, Wang and colleagues fused the Dendra2 coding sequence to the 5' and 3'UTRs of sensorin mRNA to study local translation at the sensory neuron-motor neuron junction of an Aplysia cell culture system (Wang et al., 2009). Sensorin is a peptide neurotransmitter whose UTR sequences drive the mRNA to be concentrated at synapses,

thus allowing the researchers to visualize local translation specifically at the synapse. One advantage of using Dendra2 over Kaede and other similar PAFPs is that it is activated in the 488-nm laser range, well clear of the UV-activation range and thus allowing less chance of cellular toxicity. The Venus fluorescent reporter has recently been adapted to visualize the activity-dependent dendritic localization of PSD-95 (Tatavarty et al., 2012; Ifrim et al., 2015). Ifrim et al. developed a unique fusion construct in which the Venus reporter was inserted after the 5'UTR of the PSD95 mRNA sequence. Used in this way, the Venus reporter (1) fluoresces just prior to and during translation of the PSD95 open reading frame in real-time and (2) visualizes de novo protein synthesis of PSD95 specifically in the dendritic spine (Ifrim et al., 2015). Finally, another approach to monitoring local translation is the use of biarsenical fluorescent dyes such as FlAsH and ReAsH (Ju et al., 2004; Martin et al., 2005; Rodriguez et al., 2006).

TimeSTAMP, Luciferase flash kinetics and, MiniSOG

TimeSTAMP is another method that can be used to monitor new protein synthesis that does not rely on either photoconversion or chemical tags (Lin et al., 2008; Lin and Tsien, 2010). An epitope tag is attached to a protein of interest which can be removed in the presence of a specific protease. Introduction of a drug prevents this process from occurring thus allowing newly synthesized proteins to retain their tags. TimeSTAMP allows the surveying of the whole brain in living animals.

Recently, a luciferase-based approach that utilizes flash kinetics has also been developed (Na et al., 2016). This approach uses a small luciferase protein derived from *G. princeps* whose signal is dependent on the presence of coelenteraizine. Luciferase is fused to the protein of interest so that, upon translation, luciferase will react with coelenteraizine to produce a light signal. The signal decays rapidly (< 9s for the tested Arc-luciferase

construct) so translating RNAs can be unambiguously detected even if coelenteraizine is still present. Thus, flashes of light signal indicate de novo protein synthesis of the protein of interest. Importantly, this technique utilizes wide-field microscopy, whereas other reporter systems usually require confocal or 2-photon imaging to detect signals above photobleaching (Na et al., 2016), making this assay accessible to a greater number of labs. Moreover, the authors who developed this assay suggest that estimation of the number of proteins synthesized in a local translation hot spot is within reach if combined with SUnSET; however, this is yet to be verified.

There are a growing number of electron microscopy (EM) studies detecting polyribosomes near spine bases that relocate in the spine head upon memory consolidation (Ostroff et al., 2010; Ostroff et al., 2014; Ostroff et al., 2017). However, new protein detection at the EM level has been difficult due to a lack of electron dense tags useful for specific protein visualization. Shu et al. has developed a small fluorescent protein, MiniSOG (mini singlet oxygen generator), that can be fused to a protein of interest and expressed in cells, tissue, and living organisms and detected both at the light level and by EM (Shu et al., 2011). Fluorescence photooxidation of DAB can be achieved using fluorescence and 1O2 from MiniSOG. It's been suggested that spatiotemporal control of local photogeneration of 1O2 will rapidly inactivate proteins of interest (Shu et al., 2011), perhaps allowing one to detect site specific translation at the light and subsequently at the EM level. Although, techniques like miniSOG hold promise, detection of specific translation at the EM level is still lacking.

Ensuring that synthesis of the protein of interest occurs in dendrites and not the soma

An additional requirement for local protein synthesis assays is to demonstrate that the synthesis occurs in the dendrites and not in the soma. In a perfect world, researchers

would sever dendrites and test for the appearance of newly synthesized proteins within isolated dendritic compartments. While Martin et al., have successfully demonstrated local synthesis in processes in the absence of a soma in cultured Aplysia neurons, it has proven far more difficult in mammalian neurons (Martin et al., 1997). Several labs have introduced microlesions to separate the dendrite and the soma in acute hippocampal slices (Kang and Schuman, 1996; Ouyang et al., 1999; Waung et al., 2008). Kang and Schuman used this approach to show the local synthesis of BDNF occurred in the dendrites with BDNF-mediated potentiation (Kang and Schuman, 1996). While this approach appears optimal, it does have its caveats. We have noted that severing dendrites in this way can induce local protein synthesis of some proteins, often confounding the results when assessing activity-dependent local protein synthesis (*unpublished observation*). Thus, methods have evolved that have adapted time-lapse imaging of fluorescent proteins fused to specific dendritic targeting sequences to measure de novo protein synthesis. Such measures as continuously photobleaching the soma have been used to avoid signal from the soma confounding dendritically translated fluorescent proteins (Aakalu et al., 2001). While a significant improvement over previous methods, continuous photobleaching can be toxic to the cell.

The use of photoactivated fluorescent reporters, such as Kaede, does not require constant photoconversion, in most cases, to specifically detect protein synthesis in dendrites. However, these experiments require careful controls to be able to draw this conclusion. As observed by many laboratories, locally synthesized protein synthesis occurs in hot spots of the dendrites (Aakalu et al., 2001; Job and Eberwine, 2001; Kim et al., 2013b; Sosanya et al., 2015b). Therefore, measuring the rate of diffusion of the protein of interest by photoconverting part of the dendrite and measuring diffusion over time (i.e. does the photoconverted protein diffuse into the unconverted dendrite) allows one to measure diffusion rates. For example, Raab-Graham and colleagues used such an approach

by fusing Kaede to the mRNA of Kv1.1 plus its 3'UTR. It was found that Kaede-Kv1.1 did not diffuse, and stayed localized to hot spots within the dendrite (Raab-Graham et al., 2006). Thus, proteins presumably function close to the site of synthesis and are easily detected in dendrites when fused to photoactivatable proteins.

Additional approaches include only activating specific synapses and measuring the increase in protein expression within the activated part of the dendrite. For example, Huber and colleagues measured Arc expression after local perfusion of the mGluR agonist DHPG in distal dendrites (Waung et al., 2008). This technique relies on the fact that Arc expression occurred only in the region of mGluR activation and that protein made in the soma cannot traffic to the distal dendrites within the limited time that new Arc protein expression was detected.

Finally, one of the more thorough experimental approaches that allows one to test for function of dendritically localized mRNAs is the generation of a mouse where the dendritic targeting sequence of the protein of interest is knocked out. Mayford and colleagues created a specific knockout deleting only the dendritic targeting sequence in the 3'UTR of CaMKIIα. Importantly, ISH demonstrated that CaMKIIα mRNA was absent from the dendritic fields while still present in the soma of the CA1 hippocampal neurons in the knockout mouse relative to the wildtype litter mate. Moreover, knockout mice showed deficits in late long-term potentiation (L-LTP), a cellular model for learning and memory and at the behavioral level during a memory consolidation test (Miller et al., 2002). While these data are impressive and striking, recent data suggesting that RNA-binding proteins that compete for mRNAs may confound the interpretation of these results (Sosanya et al., 2013). The removal of a dendritic targeting sequence within the 3'UTR containing RBP motifs could possibly free up RBPs to bind to lower affinity target mRNAs which may affect synaptic plasticity and/or behavior. Still, with the development of tools

like CRISPR, where making knockouts are quicker and more cost effective, site-specific knockout of proteins will facilitate the physiological relevance of localized translation of specific targets.

HOW HIGH THROUGHPUT ASSAYS MAY ANSWER LONG-STANDING QUESTIONS IN THE FIELD OF LEARNING AND MEMORY BY IDENTIFYING NOVEL PROTEINS INVOLVED IN SYNAPTIC PLASTICITY

The usage of high-throughput experiments has the potential to point researchers into new directions to answer many long-standing questions in the field of learning and memory. As mentioned above, Niere et al., recently showed that 75% of the PSD changes in protein expression within an hour of mTOR inhibition in vivo (Niere et al., 2016). Although this data represents the overall population of dynamic protein changes, individual proteins from this data can be assayed to determine site-specific localization of their mRNAs or local synthesis. Moreover, predictions can be made to identify putative translation regulatory factors such as microRNAs and RBPs based on their mRNA sequences. Figure 1.6 outlines a few recent publications that address important questions in neuroscience. Here we provide examples of how others address these questions to assist in RNAseq and proteomic follow up studies.

Synaptic tagging and capture

The synaptic tagging and capture hypothesis (Frey and Morris, 1997; Redondo and Morris, 2011) proposes that synapses undergoing long-term changes to synaptic efficacy are somehow “tagged” in a way that allows localized proteins in one synapse to be “captured” from a protein pool to activate nearby synapses. Sosanya et al. has shown that mTOR may serve as one such tag (Sosanya et al., 2015b). They demonstrated that dendritic

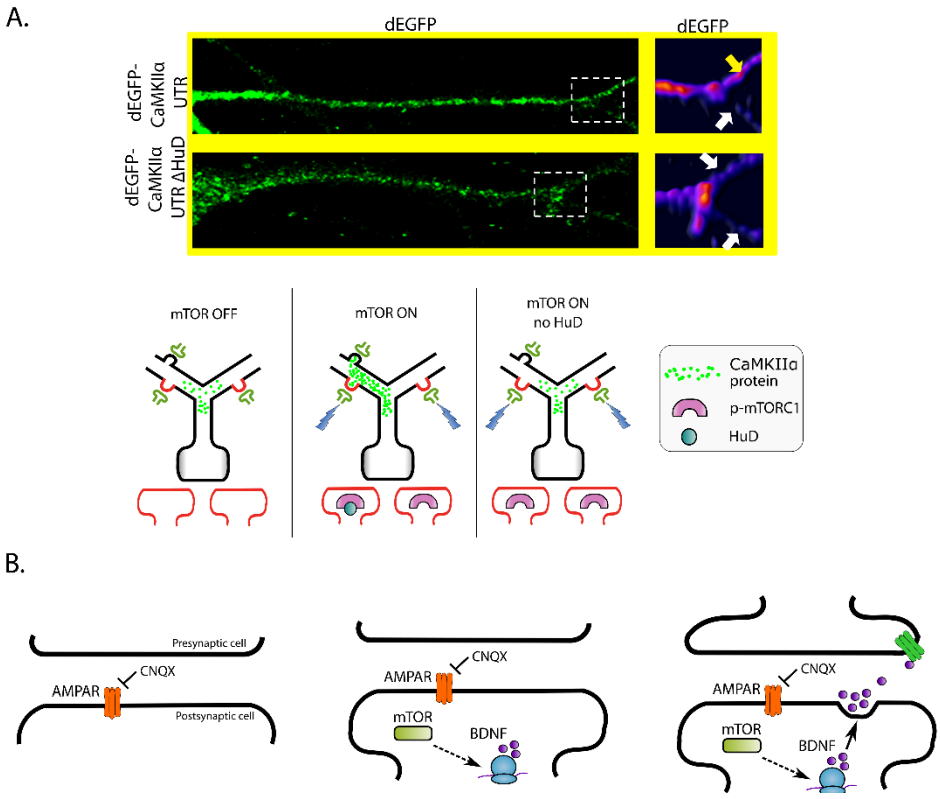


Figure 1.6: Approaches to functional validation of high-throughput experiments

A. Example of characterization of locally translated mRNAs involved in synaptic tagging and capture hypothesis. RNA sequencing of RNA-binding protein targets may show site specific expression based on the localized trafficking of the RNA-binding protein. (*Top*) For example, HuD target CaMKIIα mRNA selectively targets one dendritic branch over the other requiring HuD binding, as identified by the translational reporter construct where the coding sequence of dEGFP is fused to the 3'UTR of CaMKIIα mRNA. Deletion of the HuD binding site in the 3'UTR prevents CaMKIIα from entering into the branches and accumulates at the branch point as indicated in the branches pseudo-colored as a measure of intensity. Figure from Sosanya et al., 2015. (*Bottom*) From these data the working model of HuD's role in the synaptic tagging and capture hypothesis is illustrated. Panels represent different states of a neuron and dendritic branch point. Protrusions from top branches indicate dendritic spines. Green shapes adjacent to the spines represent presynaptic neurons. Dendritic spines of interest (red) are shown in further detail below each image of the neuron within a panel. (*mTOR OFF)* With no signal, CaMKIIα translocation does not exhibit branch preference. (*mTOR ON*) When mTOR is active, HuD targets its mRNAs into the tagged synapse. Presynaptic signals (blue lightning bolts) trigger translocation of CaMKIIα mRNA to an active synapse. (*mTOR ON or no HuD*) In the absence of HuD or mTOR, CaMKIIα does not show branch preference. **B.** Model of transsynaptic signaling based on data from (Henry et al., 2012). CNQX-mediated inhibition of AMPA receptors resulting in mTOR activation. mTOR-activation-dependent BDNF synthesis then results in BDNF secretion and binding to presynaptic terminals.

branch-specific expression of the CaMKIIα mRNA is mediated by the RNA binding protein HuD. mTOR stabilizes the mRNA and prevents degradation. Thus, mTOR serves as the tag while HuD captures CaMKIIα mRNA to promote its branch specific expression (Figure 1.6A). Data provided by high-throughput experiments that map RNA binding sites such as CLIP-SEQ used in concert with protein visualization techniques may reveal other HuD target mRNAs that have branch specific expression.

Trans-synaptic Signaling

An important but often overlooked feature involved in neuronal homeostasis is trans-synaptic signaling, the process by which proteins secreted from the postsynaptic spine can trigger retrograde activation of presynaptic neurons. Recently, it was shown that mTORC1 activation leads to the expression of BDNF which acts as a retrograde messenger to stimulate further neurotransmitter release from the presynaptic terminal (Henry et al., 2012) (Figure 1.6B). A more thorough construction of pre- and postsynaptic signaling/regulatory networks through high-throughput experiments will be invaluable in determining the nature of this postsynaptic engagement.

MOTIVATION FOR STUDIES

The basis of many diseases and neuronal dysfunctionalities arise from defects in synaptic plasticity. The molecular substructure that underlies synaptic plasticity, in part mediated by the protein kinase mTOR, forms the basis of learning and memory systems within the brain. While numerous strides have been made in recent decades with regards to the molecular composition of learning and memory systems, the application of high-throughput technologies has the potential to expand the field further and give researchers

a dynamic look at the complex web of interactions on the levels of both the proteome and the transcriptome. To this end, we have undertaken separate approaches at the level of protein and RNA that apply such high-throughput methodologies to better understand the basis of neuronal diseases. The pivotal role that mTOR plays in these processes cannot be overstated. The research in the following chapters apply many of the principles reviewed in this introduction to better understand the mechanisms by which mTORC1 dysregulation leads to disease states associated with learning and memory processes at the synapse.

mTOR lies at the intersection of numerous signaling pathways at the synapse and controls local dendritic translation processes leading many researchers to speculate that it plays a key role in the basis of neuronal dysfunction (Dazert and Hall, 2011; Wang and Proud, 2011; Laplante and Sabatini, 2012; Betz and Hall, 2013; Costa-Mattioli and Monteggia, 2013; Kaeberlein, 2013; Wang et al., 2013; Abelaira et al., 2014; Lipton and Sahin, 2014; Wong, 2014; Crino, 2015; Lan et al., 2016; Sato, 2016). Furthermore, the long-standing belief in the field is that disease of overactive mTORC1 activity are a result of excessive protein synthesis (Hoeffer and Klann, 2010). The discovery that mTOR inhibition with rapamycin leads to the increase in the expression of Kv1.1 (Raab-Graham et al., 2006) and subsequent elucidation of the mechanism by which this occurs (Sosanya et al., 2013) has raised the question of whether other proteins increase in expression as well and the contribution of these proteins to neuronal disease. Thus, Chapter 2 focuses on the application of an unbiased mass spectrometry approach to detail the wide spectrum of proteins under active suppression or activation by mTORC1 and dysregulation of these proteins in diseases such Alzheimer's disease, epilepsy, and autism spectrum disorders.

Numerous bioinformatic approaches exist that may be useful to characterize the interaction between proteins, determining potential binding targets of RNA-binding proteins, and for drug discovery. In Chapter 3, we begin with the observations that 1)

seizures are commonly observed between epilepsy, AD, and ASD (Zeng et al., 2009; Noebels, 2011; Wong, 2014; Lee et al., 2015) and 2) that these three disorders are known to be a result of overactive mTOR. We then investigate if a common phenotype underlies these disorders through the generation of a protein-protein interaction network that identifies PARK7 as a unique hub connecting these diseases. We use the results of this analysis to explore the role of RBPs in neuronal dysfunction through bioinformatics and biologically validate some of our results using the tuberous sclerosis complex (TSC) mouse model which exhibits epileptic and autistic phenotypes. These findings underscore the potential for high-throughput methodologies and bioinformatics analysis to reveal underlying connections between proteins involved in neuronal disease.

Fragile-X Mental Retardation Protein (FMRP) is an RBP that regulates translation of dendritically-localized mRNAs. Mutations in the *Fmr1* gene that encodes FMRP lead to various cognitive defects such as intellectual disability and autistic phenotypes (Bassell and Warren, 2008; De Rubeis et al., 2012). Previously, our lab has identified a connection between FMRP activity and mTOR activity in animals upon pharmacological treatment with a rapidly acting antidepressant (Workman et al., 2013; Workman et al., 2015; Wolfe et al., 2016). Chapter 4 focuses on the usage of RNA immunoprecipitation and RNA sequencing to determine the targets bound to fragile-X mental retardation protein (FMRP) in the context of a rapidly acting antidepressant dependent on mTOR activity. Traditionally prescribed antidepressants are well-known for their low response rates and poor effectiveness (Turner et al., 2008; Fournier et al., 2010; Undurraga and Baldessarini, 2012; Greenberg et al., 2015). Furthermore, there the symptoms used to characterize depression are numerous and show little overlap in descriptive categories (Lohoff, 2010; Lux and Kendler, 2010; Goldberg, 2011; Hybels et al., 2011; de Vos et al., 2015; Lieblich et al., 2015; Fried, 2017). Here, we attempt to “reverse engineer” depression using a new class

of rapidly acting antidepressants to determine the RNA populations controlled by FMRP they affect upon administration.

Collectively, the results of these studies demonstrate that high-throughput assays, bioinformatics analysis, and experimental follow-up have great potential in unraveling the details of the molecular pathways underlying neuronal dysfunction in learning and memory. The data will inform future studies in drug discovery for the creation of more targeted techniques and approaches to treating neuronal disease.

Chapter 2: Rapid repression of mTORC1 shifts the regional expression of disease-related protein ensembles[3,4]

INTRODUCTION

The mechanistic/mammalian target of rapamycin complex 1 (mTORC1) is a serine/threonine protein kinase that is highly expressed in many cell types (Kim et al., 2002). In the brain, mTORC1 tightly coordinates different synaptic plasticities—long-term potentiation (LTP) and long-term depression (LTD)—the molecular correlates of learning and memory (Tang et al., 2002; Hay and Sonenberg, 2004; Hou and Klann, 2004; Bove et al., 2011). Because mTORC1 is at the core of many synaptic signaling pathways downstream of glutamate and neurotrophin receptors, many hypothesize that dysregulated mTORC1 signaling underlies cognitive deficits observed in several neurodegenerative diseases (Hou and Klann, 2004; Kelleher et al., 2004; Takei et al., 2004; Tsokas et al., 2005; Gong et al., 2006; Antion et al., 2008b; Pei and Hugon, 2008; Swiech et al., 2008; Hoeffer and Klann, 2010; Sharma et al., 2010; Ricciardi et al., 2011; Santini and Klann, 2011; Costa-Mattioli and Monteggia, 2013). For example, mTORC1 and its downstream targets are hyperactive in human brains diagnosed with Alzheimer's disease (AD) (An et al., 2003; Li et al., 2005; Yates et al., 2013). Additionally, in animal models of autism spectrum disorder (ASD), altered mTORC1 signaling contributes to the observed synaptic dysfunction and aberrant network connectivity (Zhou et al., 2009a; Sharma et al., 2010; Bateup et al., 2011; Luikart et al., 2011; Ricciardi et al., 2011; Zoghbi and Bear, 2012; Bateup et al., 2013; Gkogkas et al., 2013; Weston et al., 2014). Furthermore, epilepsy,

[3] Parts of this chapter originally published in Niere, F., Namjoshi, S., Song E., Dilly, G.A., Schoenhard, G., Zemelman, B.V., Mechref, Y., Raab-Graham, K.F. (2016). Analysis of Proteins That Rapidly Change Upon Mechanistic/Mammalian Target of Rapamycin Complex 1 (mTORC1) Repression Identifies Parkinson Protein 7 (PARK7) as a Novel Protein Aberrantly Expressed in Tuberous Sclerosis Complex (TSC). Mol Cell Proteomics 15(2): 426-44. doi: 10.1074/mcp.M115.055079. I wrote half of the manuscript, performed bioinformatics analysis, prepared all tables, and prepared figures 2.1-2.3.

[4] Some tables and figures available online only from Molecular Cellular Proteomics.

which is common in AD and ASD, has enhanced mTORC1 activity (Crino, 2008; Pun et al., 2012; Brewster et al., 2013; Wong, 2014; Sosanya et al., 2015a).

Phosphorylation of mTORC1, considered the active form, is generally regarded to promote protein synthesis (Laplante and Sabatini, 2012). Thus, many theorize that diseases with overactive mTORC1 arise from excessive protein synthesis (Hoeffer and Klann, 2010). Emerging data, however, are demonstrating that suppressing mTORC1 activation can trigger local translation in neurons (Raab-Graham et al., 2006; Sosanya et al., 2013). Pharmacological antagonism of N-methyl-D-aspartate (NMDA) receptors, a subtype of glutamate receptors that lies upstream of mTOR activation, promotes the synthesis of the voltage-gated potassium channel, Kv1.1, in dendrites (Raab-Graham et al., 2006; Sosanya et al., 2013). Consistent with these results, in models of temporal lobe epilepsy there is a reduction in the expression of voltage-gated ion channels including Kv1.1 (Poolos and Johnston, 2012; Brewster et al., 2013; Sosanya et al., 2015a).

Interestingly in a model of focal neocortical epilepsy, overexpression of Kv1.1 blocked seizure activity (Wykes et al., 2012). Since both active and inactive mTORC1 permit protein synthesis, we sought to determine the proteins whose expression is altered when mTORC1 phosphorylation is reduced in vivo.

Rapamycin is an FDA-approved, immunosuppressive drug that inhibits mTORC1 activity (Seto, 2012). We capitalized on the ability of rapamycin to reduce mTORC1 activity in vivo and the unbiased approach of mass spectrometry to identify changes in protein expression. Herein, we provide evidence that mTORC1 activation bidirectionally regulates protein expression, especially in the PSD where roughly an equal distribution of proteins dynamically appear and disappear.

EXPERIMENTAL PROCEDURES

Sample preparation

We used three sets of paired sibling male Sprague Dawley rats (seven to nine weeks old) that were housed together. Within each pair, one received rapamycin (10 mg/kg) and the other received an equal volume of DMSO (carrier, control) via intraperitoneal (i.p.) injection. After one hour the animals were sacrificed. Cortices were homogenized, nuclei were pelleted by low speed centrifugation (100 x g) and the resulting supernatant was analyzed as cell lysates (L). To obtain synaptoneurosomes, homogenized cortices were processed as described (Workman et al., 2013). An aliquot of the synaptoneurosome fraction from each animal was solubilized in 1% triton X-100 (10 min) and centrifuged (12000 x g) to yield a triton X-100-soluble fraction (soluble, S) and a triton X-100-insoluble fraction (pellet, PSD). To prepare protein for LC-MS/MS, lysates, soluble, and PSD fractions were further solubilized in SDS-sample buffer. SDS-solubilized fractions were run on a 10% SDS-polyacrylamide gel (4 min). The migration was stopped as the ladder began to separate. Gel plugs containing the sample were sectioned and sent for mass spectrometry analysis. Animal experiments were performed according to the National Institutes of Health's Guide for the Care and Use of Laboratory Animals and approved by the UT-Austin Institutional Care and Usage Committee.

Western blot analysis

Proteins were separated by SDS-polyacrylamide gel electrophoresis (PAGE). To visualize the proteins, we used the following antibodies: rabbit anti-phospho-mTOR Ser2448 (1:2,000; Cell Signaling), mouse anti-mTOR (1:5,000; Life Technologies), mouse anti-PSD-95 (1:10,000; NeuroMab), mouse anti-synapsin 1 (1:10,000; Synaptic Systems),

rabbit polyclonal anti-α tubulin (1:50,000; Abcam), mouse ribosomal S6 (1:1000, Cell Signaling) and rabbit anti-phospho-S6 (1:1000, Cell Signaling). Membranes were subsequently incubated in fluorescence-conjugated secondary antibodies (AF680, Life Technologies; AF800, LiCor; 1:5000). Using the Odyssey CLx infrared imaging system, we obtained fluorescent images of the membranes. ImageJ (National Institutes of Health) software was used for densitometry analyses of proteins.

In-gel tryptic digestion

The gels were cut into 1mm cubes after washing gels with water for 15min. Gel pieces were transferred to a clean Eppendorf tube followed by adding 100 µl of water and incubating for 15 min. All the water was removed. This washing step was repeated with 50% acetonitrile (ACN) and 100% ACN two times. A 100 µl of 100 mM ammonium bicarbonate buffer was added followed by a 5min incubation. All the liquid was removed and the samples were subjected to drying in speed-vacuum until they were completely dried. For reduction/alkylation, 100 µl of 10 mM dithiothreitol (DTT) suspended in 100 mM ammonium bicarbonate buffer was added to the dried samples. This was followed by incubation in a water bath at 55°C for 45 min. After removing the alkylation solution, 100 µl of 55 mM iodoacetamide (IAA) suspended in 100 mM ammonium bicarbonate buffer was added into the reduced samples followed by incubating at 37.5°C for 30 min. The alkylated samples were then washed with 100 µl of 100 mM ammonium bicarbonate buffer and incubated for 5min. A 100 µl of acetonitrile was added to make one to one ratio of solutions with incubation for 15min. All the solutions were removed. The samples were then dried before digestion. Trypsin digestion solution was prepared with 4 µl trypsin in 96 µl of 100 mM ammonium bicarbonate buffer. The amount of trypsin was determined

using enzyme/protein ratio of 1:50 w/w since 200 µg of proteins were loaded into the gel. A 100 µl of trypsin digestion solution was added to the alkylated samples and then incubated for 45 min on ice. This allows the gels to absorb the trypsin digestion solution. The solution was removed if it was excessive. The samples were incubated at 37.5°C overnight. The addition of 2% trifluoroacetic acid (TFA) acidified the digestion. The tryptic digests were extracted from the gel by adding 0.1% TFA and incubating in ice water bath with sonication. The solutions were then collected in a separate clean Eppendorf tube. This extraction step was repeated with 30% ACN/0.1% TFA and 60% ACN/0.1% TFA three times. The collected peptides were then dried in a speed-vacuum and re-suspended in 0.1% formic acid prior to LC-MS/MS analysis.

Mass spectrometry

LC-MS/MS. Trypsin digested samples were subjected to LC-MS/MS analysis using Dionex 3000 Ultimate nano-LC system (Dionex, Sunnyvale, CA) interfaced to LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, San Jose, CA) equipped with a nano-ESI source. The samples were initially online-purified using a PepMap 100 C18 cartridge (3 µm, 100Å, Dionex). The purified peptides were then separated using a PepMap 100 C18 capillary column (75 µm id x 150 mm, 2 µm, 100Å, Dionex). The separation of peptides was achieved at 350 nl/min flow rate, using the following gradient: 0-10 min 5% solvent B (98% ACN with 0.1% formic acid), 10-65 min ramping solvent B 5-20%, 55-90 min ramping solvent B 20-30%, 90-105 min ramping solvent B 30-50%, 105-110min maintaining solvent B at 80%, 110-111 min decreasing solvent B 80-5%, and 111-120 min sustaining solvent B at 5%. Solvent A was a 2% ACN aqueous solution containing 0.1 % formic acid. The separation and scan time was set to 120 min. The LTQ Orbitrap Velos

mass spectrometer was operated with three scan events. The first scan event was a full MS scan of 380-2000 m/z at a mass resolution of 15,000. The second scan event was CID MS/MS of parent ions selected from the first scan event with an isolation width of 3.0 m/z, a normalized collision energy (CE) of 35%, and an activation Q value of 0.250. The third scan event was set to acquire HCD MS/MS of the parent ions selected from the first scan event. The isolation width of HCD experiment was set to 3.0 m/z while the normalized CE was set to 45% with an activation time of 0.1 ms. The CID and HCD MS/MS were performed on the 5 most intense ions observed in the MS scan event.

Data processing

Quantitation was attained employing normalized spectral counts that were calculated by Scaffold Q+ (Proteome Software, Inc., Portland, OR). The identification of proteins/peptides was achieved using MASCOT database (Brosch et al., 2009). Proteome Discoverer version 1.2 software (Thermo Scientific, San Jose, CA) was used to generate a mascot generic format file (*.mgf) which was subsequently employed for database searching using MASCOT version 2.3.2 (Matrix Science Inc., Boston, MA) Parent ions were selected from a mass range of 350-10000Da with a minimum peak count of 1. The parameters from Mascot Daemon were set to search against the UniProt Rattus database (UniProt release 2013_11). Oxidation of methionine was set as a variable modification while carbamidomethylation of cysteine was as a fixed modification. The formation of propionamide adducts on cysteine, N-terminus and C-terminus of peptides was added as variable modification due to the use of polyacrylamide gels. Trypsin was selected with missed cleavages up to two. Peptides were searched with a precursor ion mass tolerance of 6ppm or better and fragment ion mass tolerance of 1.5 Da. The MASCOT results were

imported to Scaffold version 3.6.3 (Searle, 2010) (Proteome Software, Inc., Portland, OR). Scaffold probabilistically validates the identification of peptides and proteins assigned by MASCOT using PeptideProphet (Keller et al., 2002) and ProteinProphet (Nesvizhskii et al., 2003) algorithms, respectively. Peptide identifications were accepted with a MASCOT ion score greater than 20, while protein identifications were accepted with a probability greater than 99%. The accepted proteins contain at least two identified peptides. The number of entries in the database (or subset of databases) searched was 7875 entries. The false discovery rate (FDR) was 0.2% for identification of peptides and proteins. Quantitation of identified proteins was then achieved based on spectral counts after normalization (Table S1A-D, available online). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org) via the PRIDE partner repository (Vizcaino et al., 2014; Vizcaino et al., 2016) with the dataset identifier PXD002724.

Bioinformatics

Data preparation

From the three biological replicates that were generated for each condition or treatment, namely DMSO (control) and rapamycin (Table S1A, available online), we thresholded the identified peptides by removing proteins that returned zero spectral counts in two biological replicates from each condition. We then averaged each protein across replicates for each condition, excluding zero values from the average. The thresholded protein list for each condition was used for gene ontology classification as detailed below. The data were further processed by dividing the averaged rapamycin spectral count data by the averaged DMSO spectral count data for each subcellular fraction. This produced

three lists of rapamycin fold-change over control for each of the subcellular fractions. Finally, we applied a $\log_2$ transformation to the fold-change data allowing us to easily compare difference in protein abundance for each protein found in its respective fraction. Mitochondrial proteins were filtered from fractions based on the MitoMiner database for *R. Norvegicus* mitochondrial proteins (Smith et al., 2012).

Gene ontology clustering

The unique protein lists for each condition described above were loaded into DAVID—an online tool for functional annotation and gene ontology clustering (Huang da et al., 2009a; 2009b). Data were analyzed in DAVID using the biological process (BP_FAT) ontology as well as the KEGG pathway database (KEGG_PATHWAY). We utilized the DAVID Functional Annotation Chart and Functional Annotation Clustering using the default settings (EASE score threshold = 0.1). DAVID EASE score is a modified version of the Fischer Exact p-value that allows for comparison of cluster enrichment.

Network analysis

Gene ontology clustering was first performed in DAVID to produce a Functional Annotation Chart. This chart was then loaded into the Cytoscape plugin Enrichment Map which performed further enrichment to identify clusters of high overlap and visualization in the form of an undirected network (Merico et al., 2010). We used the following Enrichment Map tuning parameters to enrich DAVID results: p-value cut-off = 0.001, Q-value cut-off = 0.05, overlap coefficient cut-off = 0.6. We utilized both Cytoscape 2.8.3 and Cytoscape 3.1.0 for our analysis (Shannon et al., 2003). Network statistics, random

network generation, and network merge analysis (to identify unique GO terms in each treatment) were performed in Cytoscape 2.8.3 using the base plugins.

Gene disease databases

For all gene lists, we matched genes in the database overlapping with our own protein list. The ASD gene list was obtained from SFARI (Simons Foundation Autism Research Initiative) (Abrahams et al., 2013). The epilepsy gene list was obtained from the CarpeDB online epilepsy gene database (http://www.carpedb.ua.edu/). The Alzheimer's disease gene list was obtained from the AlzGene database top 10 list in combination with the Alzheimer's disease pathway from the KEGG Pathway database (hsa05010) (Bertram et al., 2007; Kanehisa et al., 2014).

Statistics, data analysis, and image processing

Data processing and preparation were performed in Microsoft Excel 2010 and the R programming language (http://www.r-project.org/). Statistics and visualization was performed with the R programming language and GraphPad Prism 6. The following R packages were used for analysis: equivalence (http://cran.r-project.org/web/packages/equivalence/index.html), ggplot2, gplots (http://cran.r-project.org/web/packages/gplots/index.html), moments (http://cran.r-project.org/web/packages/moments/index.html), plyr, and reshape2 (Wickham, 2007; 2009; 2011; RC, 2014). R Scripts for generating figures and statistical analysis deposited at: https://github.com/snamjoshi/MCP2015_Niere_Namjoshi_etal. See also, Appendix C.

Bioorthogonal non-canonical amino acid tagging and Proximity ligation assay tandem (BONCAT-PLA)

Visualization of new proteins was conducted similar to (Workman et al., 2015). Following incubation of 21 DIV hippocampal neurons in 500 μL methionine-free media for 30 minutes at 37°C, azidohomoalanine (AHA, Life Technologies) and equal volume of rapamycin (final concentration = 200 nM) or DMSO were added to the media. 1 hour after treatment, cells were fixed as described by (tom Dieck et al., 2015; Workman et al., 2015). Click-it metabolic labeling was then performed according to manufacturer's instructions (Life Technologies). Biotin-alkyne was used to detect new proteins. PLA was subsequently conducted according to manufacturer's directions (Duolink, Sigma). Mouse, anti-biotin (1:500, Abcam), rabbit anti-GAP-43 (1:500, Abcam), rabbit anti-SNAP-25 (1:100, Abcam), rabbit anti-biotin (1:500, Sigma), and chicken anti-MAP2 (1:2000, Abcam) antibodies were used.

Immunocytochemistry (ICC)

Dissociated hippocampal neurons, 18-25 days in vitro (DIV), were prepared from embryonic rat pups (E17-18) similar to (Sosanya et al., 2013). For Tsc1 cultures, neurons were prepared from postnatal (0 to 3 days) Tsc1 conditional knockout pups similar to (Niere et al., 2012). Cre-2A-GFP (Puro), CMV-lentivirus (GenTarget, Inc.) was used to reduce TSC1 expression. GFP (CMV)-Puro lentiviral particles was used as a control. Virus was added to cultures at 7-9 DIV and allowed to express for 9-14 days. GFP immunofluorescence was used to identify infected pyramidal neurons. Cells were fixed in 4% paraformaldehyde for 10 minutes at 37°C and permeabilized in 0.2% Triton X-100 for 10 minutes similar to (Niere et al., 2012). Fixed cells were incubated in primary rabbit anti-GAP43 (1:500, Abcam), rabbit anti-SNAP25 (1:100, Abcam), and chicken anti-MAP2

(1:2000, Abcam). Goat anti-Synapsin (1:500, Santa Cruz) and guinea pig anti-PSD95 (1:500, Synaptic Systems) were used to visualize the pre- and postsynaptic structures respectively. Following overnight primary antibody incubation in 4°C, appropriate secondary antibodies (1:500, Life Technologies) were applied: AlexaFluor488 (AF488), AF647, AF405, and Cy3 (Jackson ImmunoResearch). Note, chicken anti-MAP2 used in conjunction with AF405 produces a granular appearance when imaged. In any given cell culture experiment, image acquisition across all conditions was performed on the same day using similar duration and intensity of excitation light. All fluorescence images were acquired on a Leica SP5 disk confocal microscope (Wetzlar, Germany) using an oil-immersion, 63X lens, sampled at 1024 X 1024.

Image analysis

For PSD-95 and synapsin colocalization analysis, Z-stacks of 9 planes at 0.5 μm/plane for each dendrite were imaged. Only one plane that had the strongest signal from the protein of interest was analyzed. The Pearson's correlation coefficient (PCC) was determined using the ImageJ plugin JACoP. The same threshold value was applied across all conditions using the threshold function in ImageJ. For BONCAT-PLA, Z-stacks of 8-9 planes at 0.5 μm/plane were imaged. For BONCAT/PLA analysis, 2 planes, for a total of 1 μm thickness that had the brightest MAP2 signals were max projected using ImageJ. The projected images were analyzed for PLA and MAP2 immunofluorescence (IF). PLA and MAP2 signals were thresholded to at least thrice above background that remained constant within an experiment. To determine the background for each signal in Metamorph, the average fluorescence intensity in a circular region of ~90 μm adjacent to the region of interest was obtained. The dendritic PLA value for each protein was normalized to the

corresponding MAP2 IF as a volume control. For BONCAT-PLA measurements, the product of the total area, intensity, and puncta number of BONCAT-PLA in each dendrite was quantified. Dendritic MAP2 IF was determined by multiplying the total area and average intensity of MAP2. The BONCAT-PLA/MAP2 values were normalized to the average BONCAT-PLA/MAP2 of control (DMSO) dendrites to reflect the total change in protein with treatment. Secondary dendrites (50 μm long and 1 μm thick) were used for analysis to minimize variability in dendritic caliber.

RESULTS

Rapid mTORC1 suppression in vivo affects synaptic processes

Having previously demonstrated that repressing mTORC1 in dissociated hippocampal neurons induced Kv1.1 mRNA translation in dendrites, we sought to identify in vivo the effects of briefly reducing mTORC1 activity on the proteomic expression of the cortex (Raab-Graham et al., 2006; Sosanya et al., 2013). We used naïve rats since our goal was to identify proteins that may be common across all diseases where mTORC1 activity is aberrantly regulated without biasing our data toward any one disease. We also took advantage of tandem mass spectrometry (LC-MS/MS) which can identify numerous proteins in a mixture and is not dependent upon access to antibody epitopes (Aebersold et al., 2013). Since local translation is rapid, we used a single intraperitoneal (i.p.) injection of rapamycin. (10 mg/kg) and harvested cortices within one hour to detect changes in protein expression that may be due to new protein synthesis (Weiler et al., 1997; Steward and Schuman, 2001; Raab-Graham et al., 2006; Waung et al., 2008; Donnelly et al., 2010; Niere et al., 2012; Ainsley et al., 2014). To verify the effectiveness of i.p. injection of rapamycin in reducing mTORC1 activity in vivo, we compared cortical lysates (Lysates)

isolated from rats injected with carrier (control/DMSO) to those injected with rapamycin by Western blotting for the active form of mTORC1 (P-mTORC1) and its downstream target ribosomal S6 (P-S6). Indeed, one hour after systemic administration, rapamycin effectively reduced the activities of mTORC1 (0.50 ± 0.14 of control) and ribosomal S6 (S6) (0.21 ± 0.08 of control), as indicated by the ratio of the phosphorylated protein over the total protein (i.e. P-mTORC1/mTORC1 and P-S6/S6; Figure 2.1A) (Sosanya et al., 2015a).

Next, we performed LC-MS/MS analyses on cortical lysates from carrier and rapamycin treated rats. To assess the consistency of spectral count measurements among replicates (i.e. 3 independent rats/treatment), we measured the average standard errors across all proteins for each treatment (Figure 2.6A; DMSO = 0.169 ± 0.006, RAPA = 0.157 ± 0.006). We also calculated spectral count correlations across all proteins for each replicate within a treatment (Figure 2.6B; $r^2 > 0.93$ (DMSO), $r^2 > 0.94$ (RAPA), average r^2 across replicates). The small variability in standard error and tight correlation among all replicates across all proteins in each treatment denote consistent sample preparation and spectral count measurements.

We first ascertained the global influence of rapidly suppressing mTORC1 by constructing biological networks using the gene ontology (GO) of each protein identified in the lysates fraction. Normalized spectral counts were averaged for each of the three biological replicates across all treatments. To find the most highly enriched biological processes, proteins were filtered into unique lists that were loaded into DAVID (Huang da et al., 2009a; 2009b). DAVID uses a modified version of Fisher's Exact Test to produce an "EASE" score that measures the significance of clustering enrichment among the proteins under consideration. Due to the high number of biological processes that resulted from DAVID clustering data, we utilized the DAVID functional annotation chart in

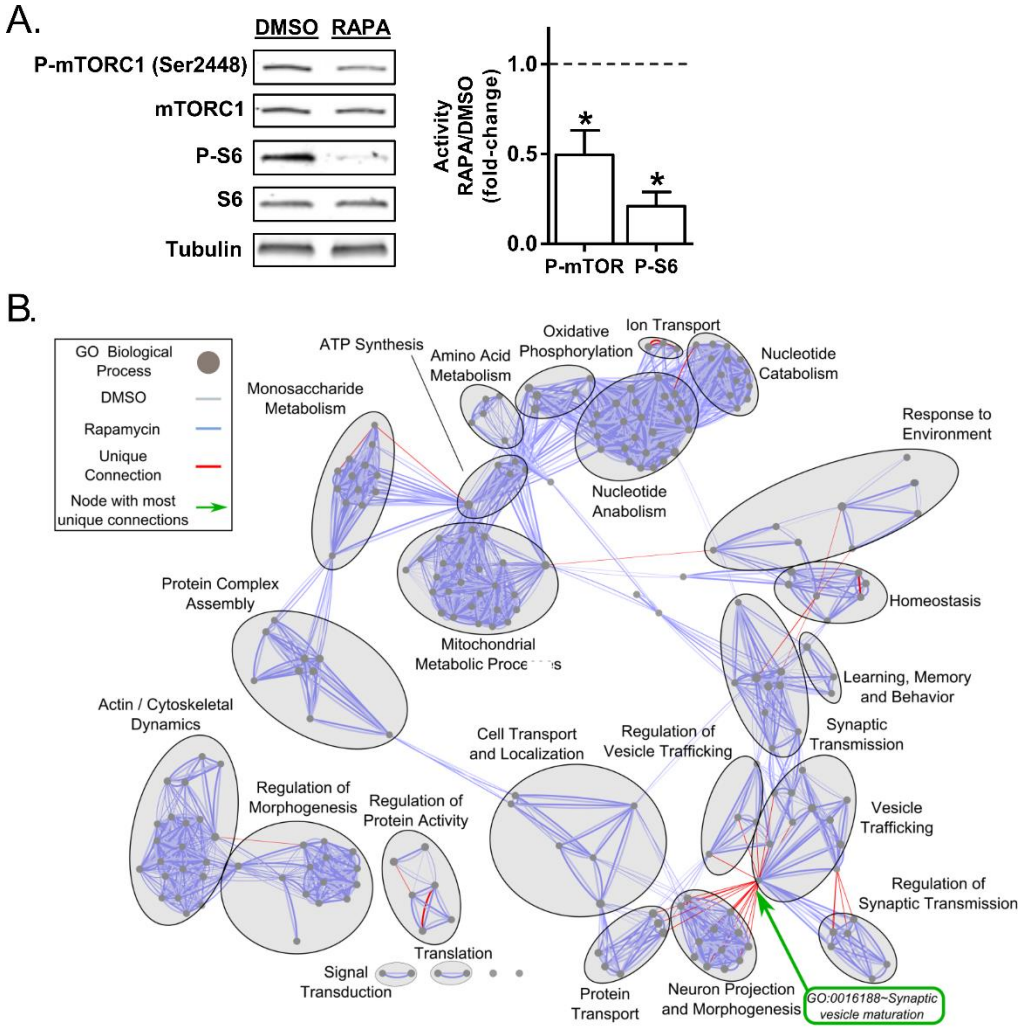


Figure 2.1: Rapamycin reduces mTORC1 activity and rapidly alters translation and synaptic function

A. One hour, systemic administration of rapamycin (10 mg/kg body weight) reduced mTORC1 activity as measured by densitometric analyses of phosphorylated mTORC1 (P-mTORC1, Ser2448) and its downstream target, phosphorylated ribosomal S6 (P-S6). To obtain fold-changes in activity, ratios of P-mTORC1 to mTORC1 and P-S6 to S6 were normalized to DMSO values. In DMSO, P-mTORC1 and P-S6 values equal 1 (dashed line, home cage control). Rapamycin, P-mTORC1 = 0.50 ± 0.14; Rapamycin, P-S6 = 0.21 ± 0.08 (Mean ± SEM). N = 3 animals per condition. Statistics: *, $p < 0.05$; Student's t test. **B.** Enrichment network comparing differences in proteins GO clustering between DMSO and rapamycin treatments for the lysate fraction. Network generated from the GO biological process database by output from the DAVID Functional Annotation Chart and further enrichment with Enrichment Map (see Experimental Procedures). Tuning parameters used to enrich DAVID results: p-value cut-off = 0.001, Q-value cut-off = 0.05, overlap coefficient cut-off = 0.6. A node represents a GO biological process containing genes from

conjunction with the Cytoscape plugin Enrichment Map to visualize gene/protein clusters with the highest overlap (Shannon et al., 2003; Merico et al., 2010; Merico et al., 2011). This approach allowed us to create biological networks organized into highly enriched functional units (Figure 2.1B). Each node in this network is a GO biological process containing a statistically significant cluster of proteins from either rapamycin or control/DMSO treatment (based on DAVID EASE score). The edge connecting the nodes illustrates the relatedness of the biological processes, such that a shorter edge indicates a greater overlap of proteins between processes. GO biological processes that contain a high overlap in common proteins were circled and given a label that generally describes the processes it contains.

To differentiate rapamycin-dependent biological processes in lysates, we overlapped control/DMSO and rapamycin networks and colored unique edges between the two fractions in red (Figure 2.1B). This approach revealed a high level of overlap between the GO biological process cluster for the DMSO (gray) and RAPA (blue) networks (Figure 2.1B). We further characterized the networks by mean closeness centrality, mean clustering coefficient, and mean neighborhood connectivity (Table 2.1) (Yu et al., 2004; Nikiforova and Willmitzer, 2007). This evaluation also indicated a stark similarity between DMSO and rapamycin networks, suggesting that acute attenuation of mTORC1 activity does not largely alter the global network. However, our analysis unexpectedly revealed that a brief

---

the experimental treatments. An edge indicates proteins shared between nodes. Edge thickness corresponds to the number of proteins shared between nodes and edge length corresponds to enrichment overlap between nodes. Light blue lines indicate DMSO treatment and dark blue indicates rapamycin treatments. Two edges of different colors connecting the same nodes indicate proteins shared between processes for both treatments. Red lines indicate unique differences in biological processes between treatments. Annotations based on general description of processes contained in the cluster. Network statistics listed in Table 2.S1.

inhibition of mTORC1 significantly affected GO biological processes observed at the synapse—ion homeostasis, metal ion transport, regulation of membrane potential, and regulation of secretion—in addition to the expected changes in translation, as indicated by the red connections (see Table 2.2 for unique GO terms between treatments). Furthermore, synaptic vesicle maturation particularly stood out with its high level of unique connections between vesicle trafficking, neuron projection and morphogenesis, and protein transport (Figure 2.1B, green arrow, lower right corner). These results suggest that acute alterations in mTORC1 activity greatly impact synaptic function.

PSD proteome is labile during brief rapamycin exposure and contains more mTORC1-responsive proteins

Because network analysis of lysates revealed stark changes in synaptic processes after one hour of mTORC1 inhibition, we isolated synaptoneurosomes from cortices of control and rapamycin-treated rats. Through a series of filtration steps and detergent solubilization, we isolated fractions of synaptoneurosomes that were soluble (supernatant, S, containing neuronal processes void of the PSD) and insoluble (pellet, P, PSD) in the detergent, triton X-100 (TX-100) (Figure 2.2A, illustration of biochemical isolation) (Fiszer and Robertis, 1967; Cohen et al., 1977). We used this modified approach for enriching PSD to allow us to identify promising proteins that directly interact with resident proteins of PSD (see Experimental Procedures). To verify the purity of the fractionation we assayed for well-characterized resident proteins of each fraction. As indicated by Western blot analysis, the pellet is enriched for PSD-95—a postsynaptic marker—and the supernatant is enriched for tubulin and synaptophysin 1—a presynaptic marker (Figure 2.2B) (Arancio et al., 1996; Biederer et al., 2002; Villasana et al., 2006).

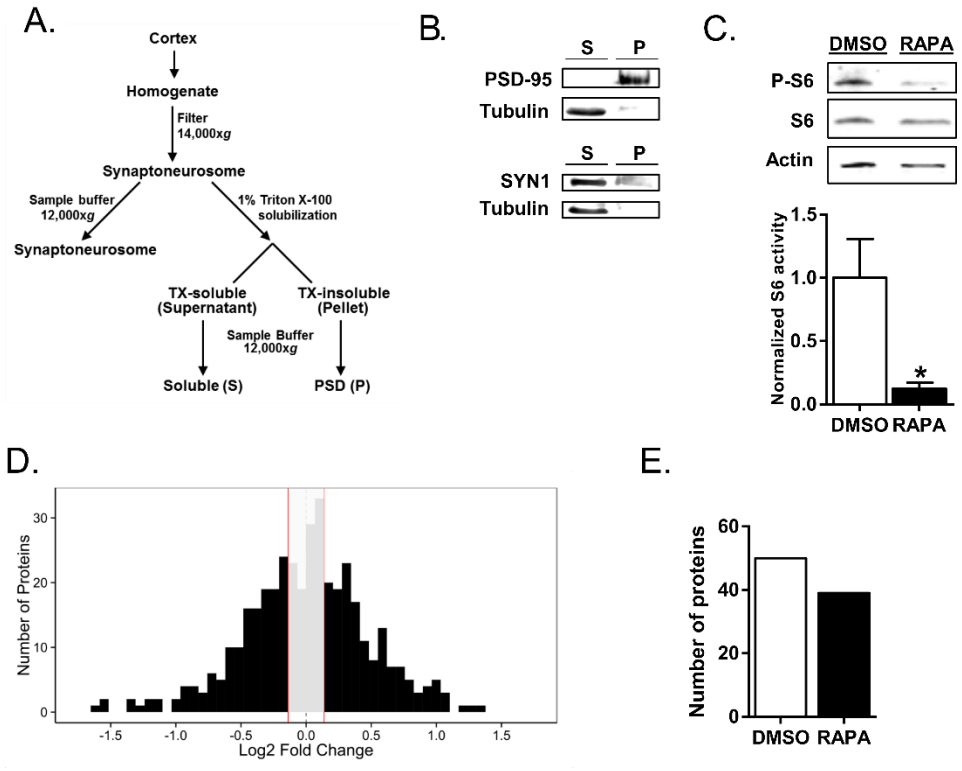


Figure 2.2: Proteins in postsynaptic region rapidly and bidirectionally respond to mTORC1 repression

A. Schematic diagram indicating the steps and reagents used to generate the PSD-enriched (P), and triton X-100-soluble (S) portions of the cortices. To isolate P and S, synaptoneurosomes were solubilized in triton-X 100 (TX-100, 1%). **B.** By Western blot analysis, TX-100-insoluble pellet (P) was enriched in postsynaptic density-95 (PSD-95, postsynaptic marker) while TX-100-soluble supernatant (S) was enriched for synapsin 1 (SYN1, presynaptic marker) and tubulin. **C.** In PSD, rapamycin reduces P-S6 level (DMSO = 1.00 ± 0.31; RAPA = 0.12 ± 0.05; N = 3 animals per condition). Actin was used as loading control since PSD fraction is void of tubulin, as shown in B. Statistics: *, $p < 0.05$; Student's t test. **D.** Histogram showing $\log_2$ transformed fold-change levels across the PSD (Figure 2.7, and Tables S4-S5, available online). Central grey box represents proteins that changed by 0-10% compared to control conditions. For PSD fraction: n = 401 proteins, $n_{\log2 \text{ fold-change} > 0} = 210$, $n_{\log2 \text{ fold-change} < 0} = 191$. Descriptive statistics: mean ± S.D. = -0.006 ± 0.5; variance = 0.25 CI [0.224, 0.281]; skewness = 0.40; kurtosis = 4.14. CI for variance calculated from X2 statistic. Bartlett's Test shows variances of lysates, PSD, and soluble fold-change distributions are unequal ($p < 0.0002$). **E.** Number of proteins in PSD where computing $\log_2$ fold-change (RAPA/DMSO) was not possible because the average spectral counts for either rapamycin or DMSO treatments was 0. These proteins are described as "out-of-range". DMSO column represents 50 proteins whose average spectral counts for the rapamycin treatment condition was 0. RAPA column represents 39 proteins whose average spectral counts for DMSO treatment condition was 0. (Table S6, available online).

Furthermore, phospho-S6 (P-S6)—a downstream marker of mTORC1 activity—was reduced in the PSD, indicating that mTORC1 was successfully inhibited by rapamycin in this fraction (DMSO = 1.00 ± 0.31; RAPA= 0.12 ± 0.05) (Figure 2.2C). Next, we subjected soluble and PSD fractions to LC-MS/MS to determine the proteins that were up- or downregulated during acute reduction of mTORC1 activity. Again, to assess consistency between replicates we measured the average standard errors and spectral count correlations across all proteins (Figure 2.7A; PSD/DMSO = 0.222 ± 0.007, PSD/RAPA = 0.198 ± 0.006; SOL/DMSO = 0.184 ± 0.005, SOL/RAPA = 0.179 ± 0.005; Figure 2.7B; $r^2 > 0.85$ (PSD/DMSO), $r^2 > 0.88$ (PSD/RAPA); $r^2 > 0.90$ (SOL/DMSO), $r^2 > 0.90$ (SOL/RAPA)).

To determine how acute mTORC1 repression affects the subcellular distribution of proteins, we examined changes in expression. The average value of each protein that was detected in rapamycin was normalized to the average value of the corresponding protein in DMSO and thus reported as the fold change of RAPA/DMSO. We found that the majority of the proteins (~73 – 75%) in PSD changed their expression profiles within one hour after reducing mTORC1 activity as determined by variance (PSD: 0.25 CI[0.224, 0.281]) and fold-changes in protein expression (see Table S4, online for PSD proteins that are insensitive to mTORC1 activity and Table S5, online for PSD proteins whose expression changes, both tables available online). Furthermore, the direction of change as indicated by the histogram is approximately equal between the expected reduction with mTORC1 inhibition (negative fold-change) and the unexpected increase (positive fold-change) in expression (Figure 2.2D). More interestingly, we discovered out-of-range proteins that could only be detected in either DMSO or rapamycin (Figure 2.2E; Table S6, available online). In PSD, 50 proteins could no longer be detected, while 39 new proteins were identified upon inhibiting mTORC1. The dynamic disappearance and emergence of out-

of-range proteins in PSD suggest that mTORC1 not only regulates the expression of proteins but also regulates them in a site-specific fashion.

Brief rapamycin exposure alters localization and expression of proteins associated with mTORC1-related disorders

Dysregulated mTORC1 signaling is implicated in many neurological disorders such as epilepsy, Alzheimer's disease (AD), and autism spectrum disorder (ASD) (Pei and Hugon, 2008; Swiech et al., 2008). We thus examined our data for disease-related proteins whose expressions were upregulated in a region-specific fashion during a brief reduction in mTORC1 activity. Like Kv1.1, these newly identified proteins that are upregulated with mTORC1 inhibition could potentially serve as novel therapeutic targets (Wykes et al., 2012). Using available databases, we compiled a list of proteins associated with epilepsy (from CarpeDB), Alzheimer's disease (AlzGene and KEGG Alzheimer's disease pathway (hsa05010)) and ASD (SFARIgene) (Table S7A, available online). We then determined which of these candidate genes appeared in our MS/MS datasets (Table S7B, available online). This list was then filtered to remove all mitochondrial genes. Proteins that were detected in two out of three samples per treatment or absent in all samples were considered for fold-expression and spatial distribution analyses as indicated on the heat maps (Figure 2.3). Notably, ~11% of the disease-related proteins showed changes where mTORC1 activity had a different impact depending on the subcellular fraction. Our analyses suggest that proteins associated with aberrant mTORC1 disorders are differentially regulated at the subcellular level by mTORC1 activity.

We detected 38 epilepsy-associated proteins (Figure 2.3, Table S8). A brief 1 hour reduction of mTORC1 signaling was sufficient to alter the expression of select proteins

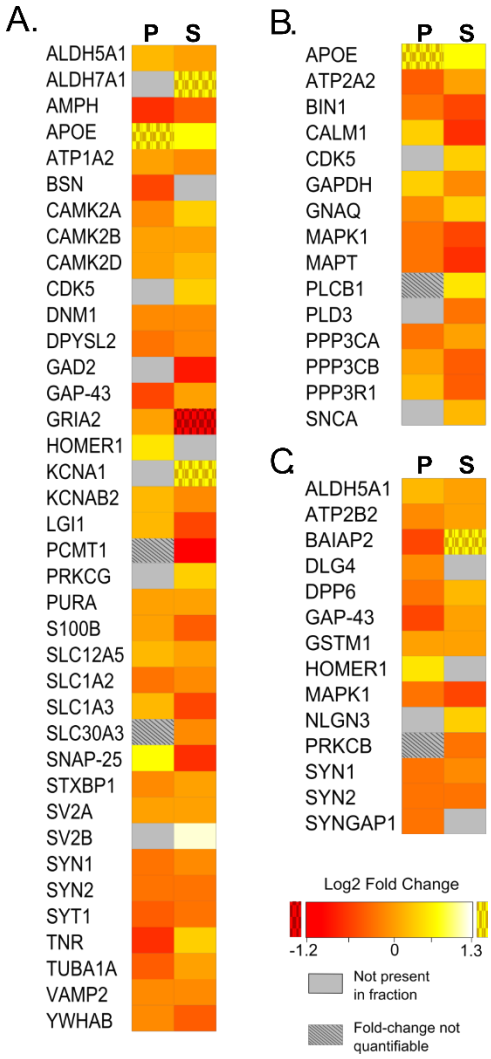


Figure 2.3: mTORC1 inhibition swiftly alters the expression of epilepsy, Alzheimer's disease (AD) and autism spectrum disorder (ASD) proteomes in a region-specific fashion

Epilepsy genes were obtained from the CarpeDB epilepsy gene database. AD genes were obtained from the AlzGene (top 10 list) and the KEGG Pathway Database for AD (KEGG hsa05010). ASD genes were obtained from the Simons Foundation Autism Research Initiative (Table S7-S8, available online). For heat maps, textured icons indicate out-of-range proteins with fold-changes of 0/Control (*red textured*) or Rapa/0 (*yellow textured*). Grey indicates that the protein was not detected in the fraction. Textured grey indicates that the fold-change was not quantifiable because only 1 of the 3 biological replicates for that treatment condition produced a non-zero spectral count value. **A.** Heatmap showing the $\log_2$ fold-change for the epilepsy proteins. **B.** Heatmap showing the $\log_2$ fold-change for the AD proteins. **C.** Heatmap showing the $\log_2$ fold-change for the ASD proteins.

bidirectionally with ~23% (red) and ~17% (yellow) increasing their expression when mTORC1 is active and inhibited, respectively (beyond ± 0.5 log_2 fold change; Table S8 online). Cyclin-dependent kinase 5 (CDK5), a molecule that regulates synaptic activity and neurological pathology, was present in Lysates and Soluble under control conditions (Figure 2.3, Table S8 online) (Barnett and Bibb, 2011; Su and Tsai, 2011). Upon repression of mTORC1, CDK5 was no longer detected in Lysates, but its expression was increased (139% of control) in Soluble. As expected, Kv1.1 protein (KCNA1) was only detected in the soluble fraction when mTORC1 was repressed, confirming our previous findings (Raab-Graham et al., 2006; Sosanya et al., 2013; Sosanya et al., 2015a). Of particular interest were the highly prominent epilepsy-associated protein leucine-rich glioma inactivated-1 (LG1) (Anderson, 2010; Yokoi et al., 2012; Kegel et al., 2013; Ohkawa et al., 2013) and the plasma membrane fusion protein synaptosomal-associated 25 kDa (SNAP-25) (Rizo and Sudhof, 2002; Catterall and Few, 2008; Selak et al., 2009; Lau et al., 2010; Ovsepian and Dolly, 2011; Tomasoni et al., 2013). Rapamycin increased the expression of LGI1 (L=159%, P=129% of control) and SNAP-25 (L=145%, P=171% of control) in Lysates (L) and PSD (P) while reducing them in Soluble (LGI1=65%, SNAP25=57% of control). These findings are very promising as LGI1, a putative tumor suppressor protein, is viewed as an antiepileptogenic protein that modulates presynaptic release probability and postsynaptic glutamate receptor function (Chernova et al., 1998; Fukata et al., 2006; Schulte et al., 2006; Zhou et al., 2009b; Fukata et al., 2010). SNAP25 mRNA is present in the dendrites and may be subject to local translation (Cajigas et al., 2012). Postsynaptically, SNAP25 regulates the trafficking of glutamate receptors to and from the membrane (Selak et al., 2009; Lau et al., 2010; Ovsepian and Dolly, 2011; Tomasoni et al., 2013). These findings are intriguing as they suggest that mTORC1

selectively and dynamically regulates regional protein expression perhaps through transport or localized translation or degradation.

We identified 15 AD-related proteins (Figure 2.3, Table S8 online). With mTOR activity, 12 proteins were similarly detected among Lysates, Soluble and PSD. Apolipoprotein E (ApoE) and α-synuclein (SNCA), hallmark proteins associated with AD and other neurodegenerative disorders, were present in Lysates and Soluble (Figure 2.3, Table S8) (Verghese et al., 2011; Korff et al., 2013). Interestingly, rapamycin shifted the spatial distribution of ApoE to PSD. We constructed a heat map to quantify in a region-specific manner the effects of mTORC1 repression on AD-related proteins (Figure 2.3, Table S8 online). We discovered that mTORC1 repression oppositely regulated the calcium-binding protein calmodulin (CALM1) in PSD and Soluble. CALM1 is hypothesized to be the central integrator of synaptic plasticity (Xia and Storm, 2005).

Since it is widely held that disrupted mTORC1 signaling strongly contributes to ASD, we examined the expression of ASD-associated proteins in our data (Kelleher and Bear, 2008; Bourgeron, 2009; Hoeffer and Klann, 2010; Sawicka and Zukin, 2012; Wang and Doering, 2013). We detected 18 proteins in control and rapamycin samples (Figure 2.3, Table S8 online). Of interest were the growth-associated protein 43 (GAP43) and Homer1b/c that reside in the PSD (Tu et al., 1999; Xiao et al., 2000; Han et al., 2013). A brief suppression of mTORC1 robustly upregulated the long Homers – Homer 1b/c – (163% of control) while downregulating GAP-43 (63% of control) (Figure 2.3, Table S8 online). We have previously shown that the mRNAs for GAP-43 and the activity-induced, short Homer, Homer1a, degrade upon mTORC1 inhibition (Workman et al., 2013). Homer1a and the long Homers 1b/c compete for the same targets (e.g. mGlu1/5 and GluN2B) within the PSD complex (Xiao et al., 1998; Shiraishi et al., 2003; Hayashi et al., 2006; Shiraishi-Yamaguchi and Furuichi, 2007). We suspect that an increase in Homer

1b/c within PSD during acute mTORC1 inhibition reduces the downstream signaling of activated receptors (e.g. mGlu1/5 and GluN2B). On the other hand, GAP-43 is closely linked to learning and memory (Benowitz and Routtenberg, 1997; Holahan et al., 2007). A report recently demonstrated that cleavage of postsynaptic Gap43 is required for NMDA-induced internalization of the α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA)-subtype glutamate receptor (Han et al., 2013). Additionally, induction of seizures by kainic acid elevated GAP-43 mRNA (McNamara and Routtenberg, 1995). Thus, disruption of GAP-43 and Homer 1b/c may serve as the initial molecular events that usher in behavioral seizures. Conversely, reducing GAP-43 and increasing Homer 1b/c may mitigate the induction or perpetuation of seizures.

Rapid mTORC1 inhibition bidirectionally regulates protein syntheses of GAP-43 and SNAP-25

A few proteins appeared to be upregulated in the PSD at one hour after mTORC1 inhibition (Figure 2.3). We focused our attention on SNAP-25 as reduction of SNAP-25 has been observed in age-related dementia including AD (Greber et al., 1999; Honer, 2003; Honer et al., 2012). This reduction has been largely attributed to a loss of synapses or dysfunction of presynaptic vesicular release (Greber et al., 1999; Washbourne et al., 2002; Honer, 2003; Honer et al., 2012). Unexpectedly, we found SNAP-25 to be upregulated in the PSD with rapamycin treatment, arguing that SNAP-25 may be important postsynaptically as well. To determine if the increase in SNAP-25 in the PSD with mTORC1 inhibition was due to new protein synthesis we used bio-orthogonal non-canonical amino acid tagging (BONCAT) and proximity ligation assay (PLA) in tandem, after repressing mTORC1 for one hour in cultured hippocampal neurons (Figure 2.4) (tom Dieck et al., 2015; Workman et al., 2015). BONCAT identifies newly synthesized proteins

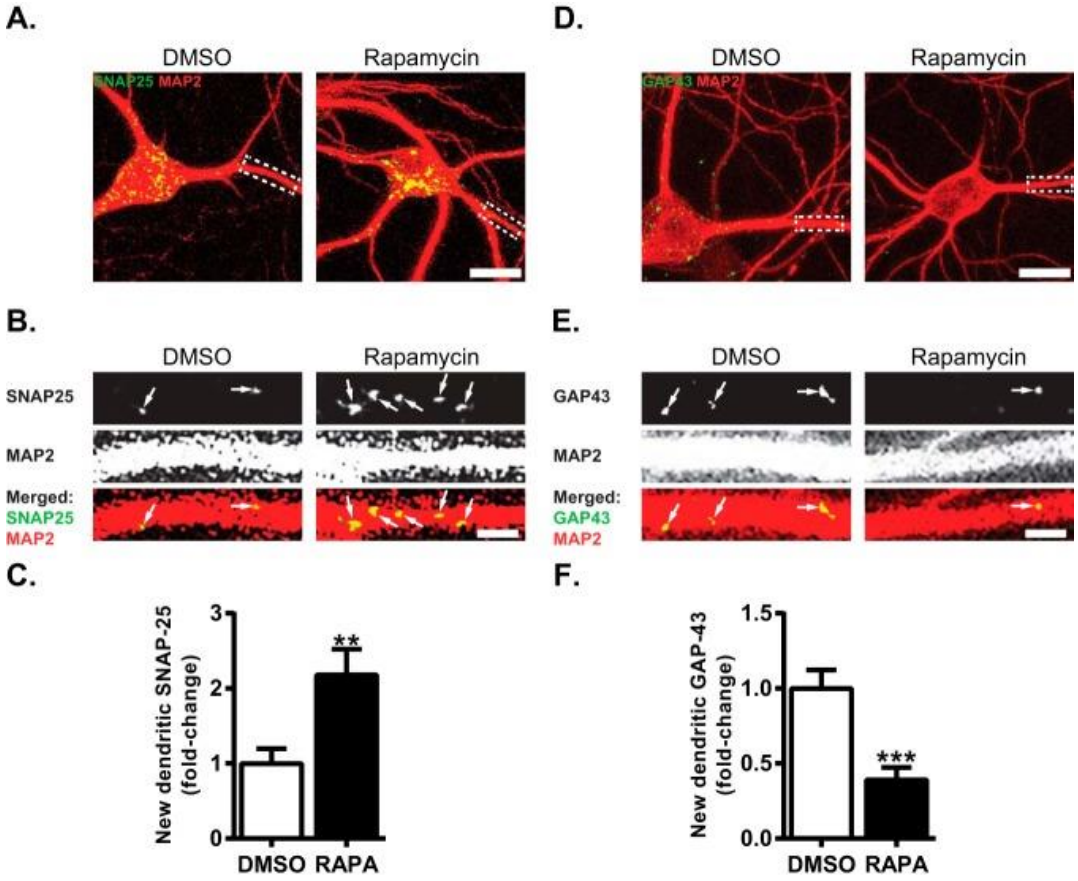


Figure 2.4: mTORC1 suppression oppositely alters the appearance of new SNAP-25 and GAP-43

A. Representative merged images of newly synthesized SNAP-25 protein (*green*) as determined by BONCAT-PLA when mTORC1 is active (*left*) or repressed (*right*) in or near MAP2-positive dendrites (*red*) of hippocampal cultured neurons. Scale bar = 20 µm. **B.** Blown up dendrite outlined in **A**. by a dotted box. Control- (*left*) and rapamycin-treated (*right*) dendrite showing new SNAP-25 (*top*), MAP2 (*middle*), and merged image (*bottom*). Arrows indicate new SNAP-25 protein. Scale bar = 5 µm. **C.** Quantification of new SNAP-25 localized within or near the dendrite. Note, SNAP-25 increases with rapamycin. (DMSO = 1.00 ± 0.20, *n* = 33 dendrites; RAPA = 2.18 ± 0.34, *n* = 37 dendrites). **D.** Representative merged images of newly synthesized GAP-43 protein (*green*) and MAP2-positive dendrites (*red*) and cell bodies treated with carrier (*left*) or rapamycin (*right*). Scale bar = 20 µm. **E.** Blown up dendrite outlined in D by a dotted box. Control- (*left*) and rapamycin-treated (*right*) dendrite showing new GAP-43 (*top*), MAP2 (*middle*), and merged image (*bottom*). Arrows indicate new GAP-43 protein. Scale bar 5 = µm. **F.** Quantification of new GAP-43 protein in or near dendrites. Note, GAP-43 protein synthesis is reduced with rapamycin. (DMSO = 1.00 ± 0.12, n = 38 dendrites; RAPA = 0.39 ± 0.08, n = 30 dendrites). MAP2 signal is saturated for visualization purposes to clearly see all dendritic branches associated with new SNAP-25 or GAP-43. Statistics: **, $p < 0.01$; ***, $p < 0.001$; Student's t test.

that incorporate an azidohomoalanine (AHA)-tagged non-canonical amino acid. Using alkyne-azide click chemistry, proteins that contain AHA are labeled with biotin. In combination with PLA, which emits a fluorescent signal only when two antibodies (anti-biotin and anti-SNAP-25) are within 30-40 nm of each other, we are able to identify new SNAP-25 protein specifically. Rapamycin, as predicted, increased new SNAP-25 in dendrites (2.18 + 0.34 of control; Figure 2.4A-C). We selected GAP-43 protein as a control that is down-regulated in PSD with mTORC1 inhibition (Han et al., 2013). Consistent with our MS/MS analyses showing reduced GAP-43 in PSD, we previously showed that Gap-43 mRNA decays with mTORC1 inhibition (Sosanya et al., 2013). Since our MS/MS data showed that GAP-43 is reduced in the PSD, we suspected that mTORC1 inhibition attenuates GAP-43 protein synthesis. Indeed, there was a significant reduction in new GAP-43 in rapamycin-treated neurons relative to control neurons (0.39 + 0.08 of control; Figure 2.4D-F), supporting our previous report that mTORC1 inhibition reduces Gap-43 mRNA (Sosanya et al., 2013). We noticed that the signal for both SNAP-25 and GAP-43 within or near dendrites was punctate—perhaps suggesting synaptic localization. We attempted to co-label pre- and postsynaptic sides of the synapse in combination with BONCAT-PLA. However, the antibodies for the presynaptic protein synapsin and the postsynaptic protein PSD-95 did not work after the BONCAT-PLA procedure. Nevertheless, these data do suggest that inhibition of mTORC1 for one hour promotes new synthesis of SNAP-25 protein while reducing synthesis of new GAP-43.

Acute mTORC1 repression bidirectionally regulates the expression of SNAP-25 and GAP-43 in the PSD

To determine if changes in expression of SNAP-25 and GAP-43 with mTORC1 inhibition are synaptic, we treated cultured hippocampal neurons with carrier (control) or rapamycin

(200 nM) for one hour. We subsequently immunostained for GAP-43 or SNAP-25, while co-staining for MAP2, PSD-95, and synapsin to visualize SNAP-25 and GAP-43 at the synapse. We then asked if the amount of SNAP-25 or GAP-43 that colocalizes with the pre- or postsynaptic terminals changes with mTORC1 activity. To do this we determined the amount of colocalized protein by plotting the fluorescence intensity of GAP-43 or SNAP-25 with the intensity of PSD-95 or synapsin and calculated the Pearson's correlation coefficient (PCC) on a pixel-by-pixel basis. Since both SNAP-25 and GAP-43 have been best characterized presynaptically, we used the PCC between SNAP-25 or GAP-43 and synapsin to normalize between conditions (Rizo and Sudhof, 2002; Powell, 2006). We found that rapamycin-treated neurons expressed more SNAP-25 that colocalized with the postsynaptic protein PSD-95, shifting its normalized PCC from 1.6 ± 0.12 to 2.1 ± 0.07 (Figure 2.5A-D). In contrast, rapamycin reduced the normalized PCC between GAP-43 and PSD-95 from 1.36 ± 0.08 to 1.0 ± 0.09 (Figure 2.5E-H). Interestingly, rapamycin had no significant effect on the normalized PCC for either SNAP-25 or GAP-43 with the presynaptic protein synapsin. Consistent with the localization of Gap-43 and Snap-25 mRNAs in the dendrites, our finding s show that GAP-43 and SNAP-25 proteins are present postsynaptically (Ovsepian and Dolly, 2011; Cajigas et al., 2012). Moreover, we demonstrate that acute mTORC1 inhibition oppositely regulates the colocalization of SNAP-25 and GAP-43 with the postsynaptic marker PSD-95.

DISCUSSION

Combined neuroproteomics approach sets a framework for new discoveries in mTORC1-dependent synapse remodeling

To generate a comprehensive view of acute mTORC1 function, we have employed tandem mass spectrometry in concert with bioinformatics analyses. Our multi-pronged

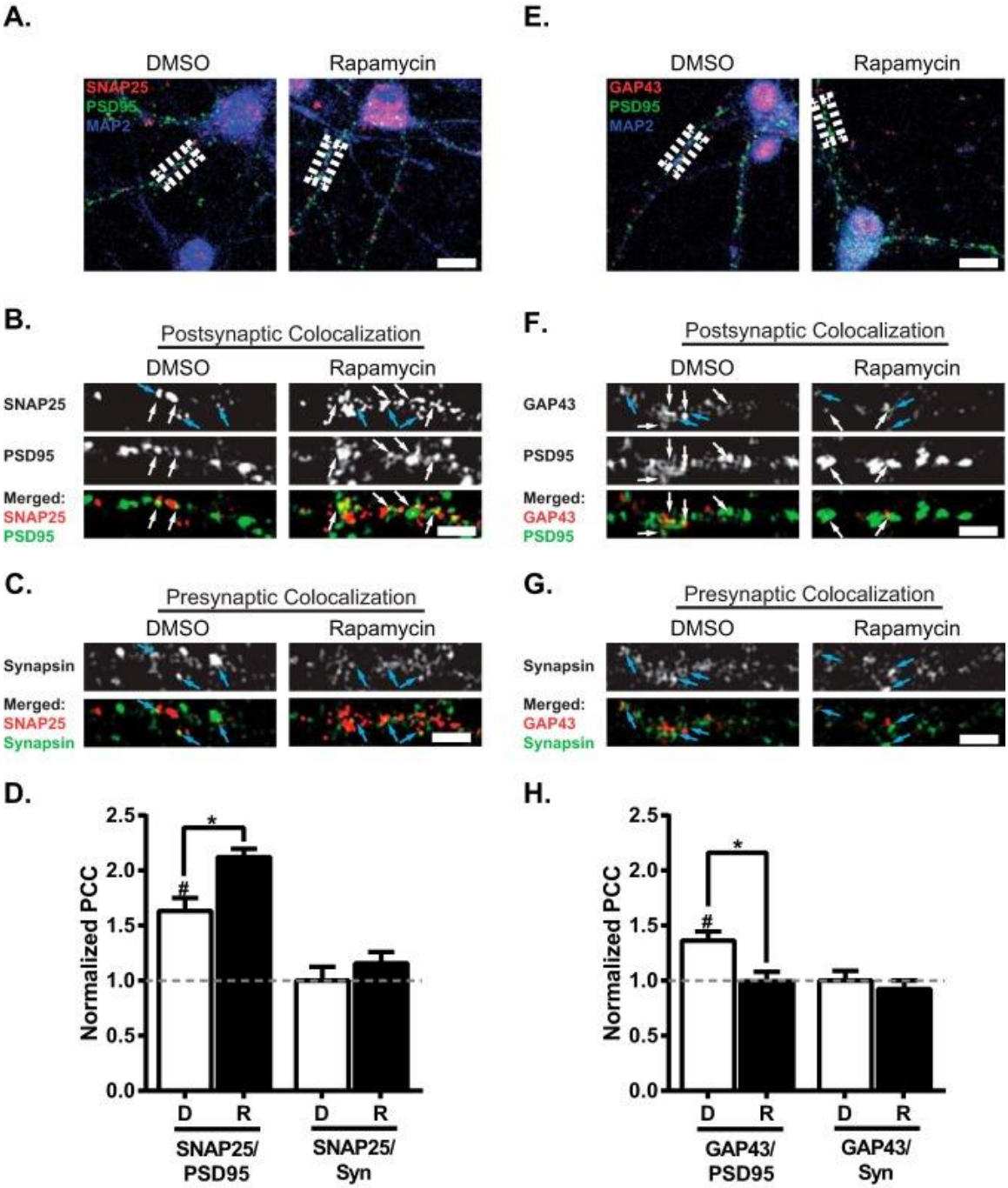


Figure 2.5: Rapid mTORC1 repression bidirectionally modifies subcellular expression of SNAP-25 and GAP-43

A. Representative images of neurons showing differential expression of SNAP-25 (*red*) colocalizing with the postsynaptic density protein PSD-95 (*green*) and the dendritic protein MAP2

approach has demonstrated that mTORC1 widely acts on protein expression by inducing differences in regional expression of proteins and the appearance and disappearance of proteins within a rapid timescale. One of the most striking findings is that ~75% of the proteins found within the PSD fraction change in abundance, suggesting that mTORC1's immediate and most pronounced function is to remodel the synapse.

(*blue*) when mTORC1 is active (DMSO, *left panel*) or repressed (Rapamycin, *right panel*). Scale bar = 20 μm. See also supplemental Figure S3. **B.** Blown up dendrites, indicated by dotted box in A, of SNAP-25 (*top*), PSD-95 (*middle*) and merged image showing colocalization (*bottom*). White arrows indicate SNAP-25 that colocalize with PSD-95. Blue arrows indicate SNAP-25 that colocalize with synapsin. **C.** Same dendrite as in **B.**, representative staining of the presynaptic protein synapsin (*top*) colocalized with SNAP-25 (*bottom, merged*). Scale bar = 5 μm. **D.** Normalized PCC averaged over many dendrites for SNAP-25 and the postsynaptic marker PSD-95 or the presynaptic marker synapsin and SNAP-25. All raw PCC values were normalized to the average PCC between SNAP-25 and synapsin in control condition (DMSO, D). Dotted line indicates the average normalized PCC between SNAP-25 and synapsin in control condition. Note, there was an increase in the PCC of SNAP-25 with PSD-95 in rapamycin treatment (DMSO, D, = 1.63 ± 0.12, n = 34 dendrites; Rapamycin, R, = 2.12 ± 0.07, n = 24 dendrites). There was no significant difference in PCC between treatments for SNAP-25 and synapsin (DMSO = 1.00 ± 0.13, n = 34 dendrites; Rapamycin = 1.16 ± 0.11, n = 24 dendrites). **E.** Representative images of neurons showing differential expression of GAP-43 (*red*), PSD-95 (*green*), and MAP2 (*blue*) when mTORC1 is active (DMSO, *left panel*) or repressed (Rapamycin, *right panel*). Scale bar 20 μm. **F.** Blown up dendrites, indicated by dotted box in **E**., of GAP-43 (*top*), PSD-95 (*middle*) and merged image showing colocalization (*bottom*). White arrows indicate GAP-43 that colocalize with PSD-95. Blue arrows indicate GAP-43 that colocalize with synapsin. See also supplemental Figure S3, available online. **G.** available online, Same dendrite as in **F.**, representative staining of synapsin (*top*) colocalized with GAP-43 (*bottom*, merged). Scale bar = 5 μm. **H.** Normalized PCC averaged over many dendrites for GAP-43 and the postsynaptic marker PSD-95 or the presynaptic marker synapsin and GAP-43. All raw PCC values were normalized to the average PCC between GAP-43 and synapsin in control condition (DMSO, D). Dotted line indicates the average normalized PCC between GAP-43 and synapsin in control condition. Note, there was a decrease in the normalized PCC of GAP-43 with PSD-95 in rapamycin treatment (DMSO, D, = 1.36 ± 0.08, n = 36 dendrites; Rapamycin, R, = 1.00 ± 0.09, n = 36 dendrites). There was no significant difference in PCC between treatments for GAP-43 and synapsin (DMSO = 1.00 ± 0.09, n = 36 dendrites; Rapamycin = 0.92 ± 0.08, n = 36 dendrites). Statistics: *, significantly different by one-way ANOVA, Tukey's multiple comparison; #, significantly different by single t test from PCC of SNAP-25 or GAP-43 with synapsin in DMSO. Chicken anti-MAP2 was imaged with 405 nm laser (see "Experimental Procedures")

mTORC1 signaling dynamically and bidirectionally regulates protein expression

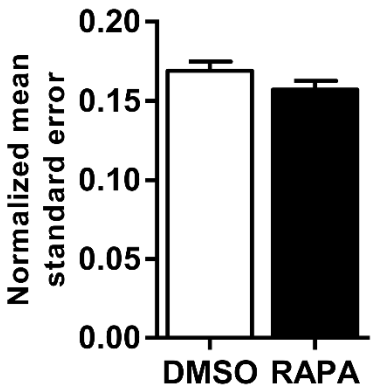
Numerous studies have established the positive function of mTORC1 activation on protein expression, but very few have examined its opposite effect (Raab-Graham et al., 2006; Auerbach et al., 2011; Laplante and Sabatini, 2012; Sosanya et al., 2013). Having previously demonstrated in cultures that repression of mTORC1 by rapamycin increases Kv1.1 translation locally in dendrites, we present here that reduction of mTORC1 activity in vivo upregulates protein expression in a region-specific manner (Raab-Graham et al., 2006). Specifically, we observe that ~30% of proteins in the PSD are elevated upon mTORC1 repression (Fig. 2.2D). One of these proteins is SNAP-25, which markedly increased in the PSD at one hour after inhibiting mTORC1. While SNAP-25, a part of the SNARE complex, has been best characterized for its presynaptic role in neurotransmitter release, reduced levels are associated with postsynaptic dysfunction. Notably, these include phenotypes that are associated with ASD such as reduced spine density, amplitude of miniature excitatory postsynaptic potentials (mEPSP), and activity induced forms of potentiation (Tavazoie et al., 2005; Aldinger et al., 2011; Durand et al., 2012; Antonucci et al., 2013; Fossati et al., 2015). Moreover, SNAP-25 has recently been shown to facilitate the removal of kainate receptors (KARs), which are implicated in epileptogenesis, from the membrane surface (Rizo and Sudhof, 2002; Catterall and Few, 2008; Selak et al., 2009; Vincent and Mulle, 2009; Lau et al., 2010; Ovsepian and Dolly, 2011; Tomasoni et al., 2013). Our results may help explain why in part rapamycin reduces seizure frequency in mTOR-related diseases, possibly through the increased expression of SNAP-25 in the PSD and the removal of KARs.

Combined mass spectrometry and bioinformatics approaches sets a framework for new discoveries in mTORC-1 dependent physiology

A great body of work has characterized the molecular hallmarks of diseases involving protracted mTORC1 dysregulation. However, the early molecular changes that contribute to the disease phenotypes have largely remained unexplored. Biochemical, immunohistochemical, and functional approaches used in conjunction can definitively determine signaling pathways. However, these approaches are time-consuming and limited to reagents available and molecules being considered. Using rapid mTORC1 repression as a paradigm of the early onset occurring in dysregulated mTORC1 signaling, we have discovered dynamic changes in protein expression globally and regionally. Our large scale proteomic approach demonstrates that the expressions of disease-associated proteins are already changing after acutely disrupting mTORC1 activity. These protein changes are maintained during the clinical presentation of many diseases in which mTORC1 is disrupted. Therefore, proteins that we have identified that change in rapid repression of mTORC1 may serve as valuable markers for the early detection of neurodegenerative illnesses. Our large-scale approach in identifying mTORC1-responsive proteins offers multiple targets that can potentially restore neurological deficits. In addition, targeting these proteins instead of mTORC1, which influences many biological processes, may serve as a more effective therapy in ameliorating mTORC1-related diseases.

SUPPLEMENTAL FIGURES[5]

A.



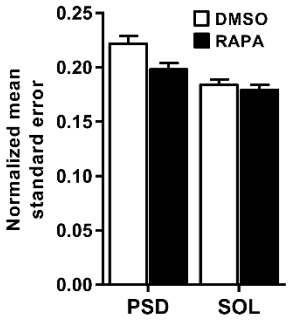
B.

Fraction	Treatment	Replicate	r^2
Lysate	DMSO	1 vs 2	0.922
		2 vs 3	0.959
		3 vs 1	0.900
	RAPA	1 vs 2	0.943
		2 vs 3	0.949
		3 vs 1	0.931

Figure 2.6: Quantitation of standard errors in spectral counts measurements across treatments for Lysates

[5] Remaining supplemental figures may be found online as supplemental material available from Molecular and Cellular Proteomics in (Niere et al., 2016). See also, Appendix E.

A.



B.

Fraction	Treatment	Replicate	r^2
PSD	DMSO	1 vs 2	0.794
		2 vs 3	0.898
		3 vs 1	0.856
	RAPA	1 vs 2	0.855
		2 vs 3	0.905
		3 vs 1	0.867
Soluble	DMSO	1 vs 2	0.832
		2 vs 3	0.900
		3 vs 1	0.907
	RAPA	1 vs 2	0.900
		2 vs 3	0.941
		3 vs 1	0.843

Figure 2.7: Quantitation of standard errors in spectral counts measurements across treatments for PSD and Soluble (SOL) fractions

A. To determine the reliability of our spectral count data, standard errors for each protein within a treatment were computed and then normalized by the mean spectral count for that protein. The y-axis represents the mean standard error across all proteins within the particular treatment and the error bars indicate the SEM for this mean value. PSD, DMSO = 0.222 ± 0.007, PSD, RAPA = 0.198 ± 0.006; SOL, DMSO = 0.184 ± 0.005, SOL, RAPA = 0.179 ± 0.005. **B.** Table showing correlations between spectral counts across different replicates for each treatment/fraction across all proteins. r^2 indicates Pearson correlation coefficient.

SUPPLEMENTAL TABLES[6]

Fraction	Treatment	Proteins	Nodes	Edges	Network Density	Characteristic Path length	Mean Closeness Centrality	Mean Clustering Coefficient	Mean Neighborhood Connectivity
Lysate	DMSO	605	224	1498	0.060	4.055	0.319 ± 0.16 $p < 0.01^{b}$	0.735 ± 0.16 $p < 0.01^{a}$ $p < 0.01^{b}$	16.458 ± 8.56 $p < 0.01^{a}$ $p < 0.01^{b}$
	Rapamycin	593	211	1370	0.062	4.139	0.325 ± 0.19 $p < 0.01^{b}$	0.739 ± 0.16 $p < 0.01^{a}$ $p < 0.01^{b}$	15.88 ± 7.32 $p < 0.01^{a}$ $p < 0.01^{b}$

Table 2.1: Summary of network statistics

Mean statistics are listed with their standard deviations. [a]Kolmogorov-Smirnov test comparing network to randomly generated network. Random networks (undirected) represent randomized forms of the networks they are compared to. [b]TOST Equivalency test comparing one network to another experimental treatment network within each fraction. Wilcoxon-Mann-Whitney test was used to determine TOST p-values.

[6] Remaining supplemental tables may be found online as supplemental material available from Molecular and Cellular Proteomics in (Niere et al., 2016). See also Appendix E

Treatment	Unique GO Terms	GO I.D.	Proteins in GO Cluster	p-value
DMSO	METAL ION TRANSPORT	GO:0030001	35	4.18×10^{-4}
	RESPONSE TO METAL ION	GO:0010038	15	5.78×10^{-4}
	REGULATION OF PROTEIN MODIFICATION PROCESS	GO:0031399	26	2.85×10^{-4}
	REGULATION OF SECRETION	GO:0051046	20	4.20×10^{-4}
	LEARNING	GO:0007612	12	1.97×10^{-5}
	NUCLEOTIDE CATABOLIC PROCESS	GO:0009166	9	8.06×10^{-4}
	NEGATIVE REGULATION OF PROTEIN MODIFICATION PROCESS	GO:0031400	16	6.92×10^{-5}
	REGULATION OF CELLULAR PROTEIN METABOLIC PROCESS	GO:0032268	35	5.91×10^{-4}
	PROTEIN HOMOOLIGOMERIZATION	GO:0051260	13	3.55×10^{-4}
	HETEROCYCLE CATABOLIC PROCESS	GO:0046700	11	8.71×10^{-4}
	SYNAPTIC VESICLE MATURATION	GO:0016188	4	5.79×10^{-4}
	NEGATIVE REGULATION OF CELLULAR PROTEIN METABOLIC PROCESS	GO:0032269	18	7.93×10^{-4}
	ION HOMEOSTASIS	GO:0050801	31	8.48×10^{-4}
	PURINE NUCLEOTIDE CATABOLIC PROCESS	GO:0006195	8	2.54×10^{-4}
	REGULATION OF MEMBRANE POTENTIAL	GO:0042391	15	8.46×10^{-4}
	TRANSLATION	GO:0006412	27	6.99×10^{-4}
	RESPONSE TO ORGANIC CYCLIC SUBSTANCE	GO:0014070	14	9.88×10^{-4}
	RESPONSE TO ORGANIC NITROGEN	GO:0010243	10	7.79×10^{-4}
Rapamycin	CELL REDOX HOMEOSTASIS	GO:0045454	10	7.19×10^{-4}
	MICROTUBULE-BASED PROCESS	GO:0007017	23	3.91×10^{-4}
	GLUCONEOGENESIS	GO:0006094	7	3.26×10^{-4}
	CELLULAR ALDEHYDE METABOLIC PROCESS	GO:0006081	8	7.77×10^{-5}
	RESPONSE TO REACTIVE OXYGEN SPECIES	GO:0000302	11	6.47×10^{-4}

Table 2.2: Unique GO Terms in Lysates by Experimental Treatment

Table shows unique GO terms between DMSO and Rapa treatments. p-values (DAVID EASE score, modified Fischer exact test p-value) generated using the DAVID Functional Annotation Chart. Average p-values for lysate fraction: 5.67×10^{-4} (DMSO), 4.32×10^{-4} (Rapa).

Chapter 3: Bioinformatic approaches identify PARK7 as a novel protein aberrantly expressed in Tuberous Sclerosis Complex[7,8]

INTRODUCTION

The process of learning and memory is dependent on the expression of a wide variety of molecules such as ion channels that can modulate the strength or weakness of synaptic transmission between neighboring neurons (Zhang and Linden, 2003; Kim and Hoffman, 2008). Since each synaptic connection is unique, a coordinated effort on the part of the molecules within the synapse is required to establish and maintain the connections between the pre- and post-synaptic cells. These processes rely on a large number of proteins, such as the mechanistic/mammalian target of rapamycin serine/threonine protein kinase (mTOR), that can alter the molecular composition of the synapse (Dazert and Hall, 2011; Wang and Proud, 2011; Laplante and Sabatini, 2012; Betz and Hall, 2013; Costa-Mattioli and Monteggia, 2013; Kaeberlein, 2013; Wang et al., 2013; Abelaira et al., 2014; Lipton and Sahin, 2014; Wong, 2014; Crino, 2015; Lan et al., 2016; Sato, 2016). mTOR and other important proteins are responsible for regulating the translation of dendritically-localized mRNA so that they can be rapid expressed in the correct temporal and spatial dimensions. RNA-binding proteins (RBPs) play a key role in these regulatory processes and disruptions result in neuronal dysfunction and disease (Giese and Mizuno, 2013; Lipton and Sahin, 2014).

---

[7] Niere, F., Namjoshi, S., Song E., Dilly, G.A., Schoenhard, G., Zemelman, B.V., Mechref, Y., Raab-Graham, K.F. (2016). Analysis of Proteins That Rapidly Change Upon Mechanistic/Mammalian Target of Rapamycin Complex 1 (mTORC1) Repression Identifies Parkinson Protein 7 (PARK7) as a Novel Protein Aberrantly Expressed in Tuberous Sclerosis Complex (TSC). Mol Cell Proteomics 15(2): 426-44. doi: 10.1074/mcp.M115.055079. I wrote half of the manuscript, performed bioinformatics analysis, prepared all tables, and prepared figures 3.1 and 3.2.

[8] Some tables and figures available online only from Molecular Cellular Proteomics.

mTOR signaling is known to be dysregulated in Tuberous Sclerosis Complex (TSC) (Huang and Manning, 2008; Swiech et al., 2008; Zeng et al., 2009). TSC is a multisystem disorder which in the central nervous system manifests itself in the form of epilepsy, ASD, and intellectual disabilities (Inoki et al., 2003; Tavazoie et al., 2005; Inoki et al., 2006; Crino, 2008; Bourgeron, 2009; Bateup et al., 2011; Zeng et al., 2011; Sahin, 2012; Bateup et al., 2013). At the cellular level, synaptic plasticity is altered in mouse models of TSC. TSC consists of two obligate subunits, TSC1 and TSC2. The TSC complex inhibits mTOR activity. Importantly, disease causing mutations in either TSC1 or TSC2 lead to overactive mTOR signaling. However, what mRNAs are repressed in TSC and how their translation is regulated are still open questions.

Parkinsonian-Associated Protein 7 (PARK7 or DJ-1) is a small 24 kDa RBP with a highly-conserved structure (Cookson, 2004) and has been shown to regulate translation of a number of RNA transcripts by direct binding (van der Brug et al., 2008). Specifically, PARK7 associates with GC-rich transcripts involved in mitochondrial homeostasis, the PTEN/Akt survival pathway, apoptosis, and glutathione synthesis. The regulatory effects of PARK7 are lost under oxidative stress conditions. Recently, it was demonstrated that yeast PARK7 homologs and TORC1 (an mTOR homolog) are co-localized in stress granules and in P-bodies (Miller-Fleming et al., 2014). Both of these structures form in the cytoplasm (in neuronal as well as other cell types) under highly stressed conditions and allow for the rapid, local translation of a specific set of mRNAs needed to quickly respond to a stressed environment or to be sequestered to prevent translation. TORC1 is also known to be localized to stress granules under heat shock (Takahara and Maeda, 2012). Finally, PARK7 binds to components of the ribosome small and large subunit as well as other proteins (Jin et al., 2007). 114 of these proteins showed significant changes in abundance upon treatment with rotenone, a pesticide that produces PD-like phenotypes in animal

models (Jin et al., 2007). These new findings strongly point to a connection between mTORC, PARK7, and translational dysfunction at the synapse.

Although PARK7 is best known for its involvement in Parkinson's disease, its role in mRNA translational regulation, synaptic plasticity, and diseases such as Autism Spectrum Disorders (ASD) or TSC, diseases associated with epilepsy and overactive mTOR signaling, is not known (Zeng et al., 2009; Noebels, 2011; Wong, 2014; Lee et al., 2015). Here we apply multiple bioinformatics approaches to better clarify the role of PARK7 in these disorders including a novel role in TSC. Our results demonstrate the power of bioinformatics in revealing unknown connections between neuronal disease and offer possibilities for how these neuronal dysfunctions may be treated on a pharmacological level through drug discovery.

EXPERIMENTAL PROCEDURES

Western blot analysis

Proteins were separated by SDS-polyacrylamide gel electrophoresis (PAGE). To visualize the proteins, we used the following antibodies: rabbit anti-phospho-mTOR Ser2448 (1:2,000; Cell Signaling), mouse anti-mTOR (1:5,000; Life Technologies), mouse anti-Park7 (1:2000; Novus Biologicals) mouse anti-PSD-95 (1:10,000; NeuroMab), mouse anti-synapsin 1 (1:10,000; Synaptic Systems), rabbit polyclonal anti-α tubulin (1:50,000; Abcam), mouse ribosomal S6 (1:1000, Cell Signaling) and rabbit anti-phospho-S6 (1:1000, Cell Signaling). Membranes were subsequently incubated in fluorescence-conjugated secondary antibodies (AF680, Life Technologies; AF800, LiCor; 1:5000). Using the Odyssey CLx infrared imaging system, we obtained fluorescent images of the membranes.

ImageJ (National Institutes of Health) software was used for densitometry analyses of proteins.

Bioorthogonal non-canonical amino acid tagging and Proximity ligation assay tandem (BONCAT-PLA)

Visualization of new proteins was conducted similar to (Workman et al., 2015). Following incubation of 21 DIV hippocampal neurons in 500 μL methionine-free media for 30 minutes at 37°C, azidohomoalanine (AHA, Life Technologies) and equal volume of rapamycin (final concentration = 200 nM) or DMSO were added to the media. 1 hour after treatment, cells were fixed as described by (tom Dieck et al., 2015; Workman et al., 2015). Click-it metabolic labeling was then performed according to manufacturer's instructions (Life Technologies). Biotin-alkyne was used to detect new proteins. PLA was subsequently conducted according to manufacturer's directions (Duolink, Sigma). Mouse, anti-biotin (1:500, Abcam), rabbit anti-GAP-43 (1:500, Abcam), rabbit anti-SNAP-25 (1:100, Abcam), mouse anti-PARK7 (1:1000, Sigma), rabbit anti-biotin (1:500, Sigma), and chicken anti-MAP2 (1:2000, Abcam) antibodies were used.

Stereotaxic in vivo virus injection and hippocampal immunohistochemistry

To remove TSC1 in Tsc1 conditional knockout mice (Tsc1tm1Djk/J, Jackson Laboratory), Cre-recombinase and GFP packaged in adeno-associated virus (AAV-Cre and AAV-GFP) were stereotaxically injected into the hippocampus of 7 to 8 week-old male mice as described (Workman et al., 2015). The following coordinates (from bregma) were used: −2.2 mm A/P, +/− 1.5 mm M/L; −2.5 mm A/P, and +/− 1.6 mm M/L. Two weeks after injection, mice were transcardially perfused with phosphate buffered saline (0.1 M)

and the brains were postfixed in 4% paraformaldehyde. 50 μm thick hippocampal slices were prepared and checked for GFP expression. Slices that expressed GFP were processed from immunohistochemistry as described (Hays et al., 2011) using the following primary antibodies: mouse anti- PARK7 (1:100, Sigma Aldrich), guinea pig anti-PSD95 (1:100, Synaptic Systems), and chicken anti-MAP2 (1:500). After overnight primary antibody incubation in 4°C, the following secondary antibodies were used: AF405, Cy3, and AF647.

Image analysis

For PSD-95 and synapsin colocalization analysis, Z-stacks of 9 planes at 0.5 μm/plane for each dendrite were imaged. Only one plane that had the strongest signal from the protein of interest was analyzed. The Pearson's correlation coefficient (PCC) was determined using the ImageJ plugin JACoP. The same threshold value was applied across all conditions using the threshold function in ImageJ. For BONCAT-PLA, Z-stacks of 8-9 planes at 0.5 μm/plane were imaged. For BONCAT/PLA analysis, 2 planes, for a total of 1 μm thickness that had the brightest MAP2 signals were max projected using ImageJ. The projected images were analyzed for PLA and MAP2 immunofluorescence (IF). PLA and MAP2 signals were thresholded to at least thrice above background that remained constant within an experiment. To determine the background for each signal in Metamorph, the average fluorescence intensity in a circular region of ~90 μm adjacent to the region of interest was obtained. The dendritic PLA value for each protein was normalized to the corresponding MAP2 IF as a volume control. For BONCAT-PLA measurements, the product of the total area, intensity, and puncta number of BONCAT-PLA in each dendrite was quantified. Dendritic MAP2 IF was determined by multiplying the total area and average intensity of MAP2. The BONCAT-PLA/MAP2 values were normalized to the

average BONCAT-PLA/MAP2 of control (DMSO) dendrites to reflect the total change in protein with treatment. Secondary dendrites (50 μm long and 1 μm thick) were used for analysis to minimize variability in dendritic caliber. To quantify dendritic PARK7 in hippocampal slices, Z-stacks of 12 planes were used at 1 μm/plane. Only 1 plane where MAP2 signal is distinct but not saturated was analyzed. For each dendrite, the product of the total area and intensity of PARK7 was determined and normalized to its corresponding MAP2 immunofluorescence as a volume control. Each dendritic PARK7/MAP2 value was normalized to the average PARK7/MAP2 of wildtype dendrites. To outline the dendrites in hippocampal slices, the MAP2 signal was saturated to define the dendrite. The "Find Edges" function in ImageJ was used to create an outline of the dendrite.

Bioinformatics

Gene ontology clustering

The unique protein lists for each condition described above were loaded into DAVID—an online tool for functional annotation and gene ontology clustering (Huang da et al., 2009a; 2009b). Data were analyzed in DAVID using the biological process (BP_FAT) ontology as well as the KEGG pathway database (KEGG_PATHWAY). We utilized the DAVID Functional Annotation Chart and Functional Annotation Clustering using the default settings (EASE score threshold = 0.1). DAVID EASE score is a modified version of the Fischer Exact p-value that allows for comparison of cluster enrichment.

Protein-protein interaction network analysis

Protein-protein interaction networks were generated using the BisoGenet plugin (Martin et al., 2010). Inputs were mapped as protein identifiers using protein-protein

interaction data from the BioGRID and MINT databases. Network was built by adding neighbors to seed proteins up to a distance of 2. Our output network was represented in terms of proteins. Networks were analyzed using degree, betweenness and closeness centrality network statistics. Hub proteins were identified using the degree network statistic.

Determining PARK7 binding motifs

A CLIP study for PARK7 revealed the presence of multiple binding motifs among the isolated targets (van der Brug et al., 2008). The authors of this work present their data as a 45-nucleotide long consensus logo that collectively summarizes the most probable nucleotides per position for all recovered CLIP tags. Since the typical RNA-binding motif is of length 6 nucleotides, we constructed a list of all possible substrings of this length from the consensus logo. Each substring was then ranked based on the sum of its probability values based on the probabilities derived from the consensus logo. We found that sequences contained the two motifs GNGCNG or CNGCNG occurred with highest probability. To find the frequency of occurrences for these motifs in a given DNA alphabet, we used the seqinr (Charif and Lobry, 2007) and Biostrings (Pages et al., 2016) packages. A graphical user interface for this process was developed using the shiny package (Chang et al., 2016) in R along with the DT package (Xie, 2016). PSD gene data obtained from the BAYES-COLLINS-MOUSE-PSD-CONSENSUS (Bayes et al., 2012) available online at G2C: Genes to Cognition database (http://www.genes2cognition.org/). All gene sequences were obtained using the biomart package (Durinck et al., 2005; Durinck et al., 2009; Smedley et al., 2015).

Connectivity Map (CMap) analysis of PARK7 PSD targets

We obtained a list of putative PSD proteins (see above) and separated the lists into upregulated or downregulated genes based on our prior mass spectrometry data (Niere et al., 2016). We imported these gene lists to the Connectivity Map software online (Lamb et al., 2006; Lamb, 2007). We took the top compounds that were returned and then used LigDig (Fuller et al., 2015) to find their pubchem IDs. The resulting IDs were then imported into R. We then used the chemmineR package (Cao et al., 2008) to determine the pubchem footprint of each compound. The similarity of these footprints can then be assessed by hierarchical clustering and visualized as a heatmap. We used the pheatmap package (Kolde, 2015) to create the clustered dendrogram and heatmap.

RESULTS

Protein-protein interaction network analysis exposes Snca and Park7 as novel protein hubs shared by epilepsy, AD and ASD

Seizures are commonly observed in epilepsy, AD, and ASD. We, therefore, considered the possibility that the proteins linked to these disorders (Figure 3.1) interact with each other in some way to elicit a shared disease phenotype. Using a Venn diagram to determine which of the previously identified candidate proteins are common in epilepsy, AD and ASD, we found that few if any were shared by the three disorders (Figure 3.1A; 2 genes from published data (*left*) and 0 genes from our data (*right*)). We next sought to determine potential protein-protein interactions (PPI) using epilepsy-, AD-, and ASD-related proteins from our MS/MS data. To do this, we first created a PPI network for each disease by seeding each network with its relevant disease-related proteins (48 for epilepsy, 44 for AD, and 19 in ASD). We then constructed each network by adding neighbors to the

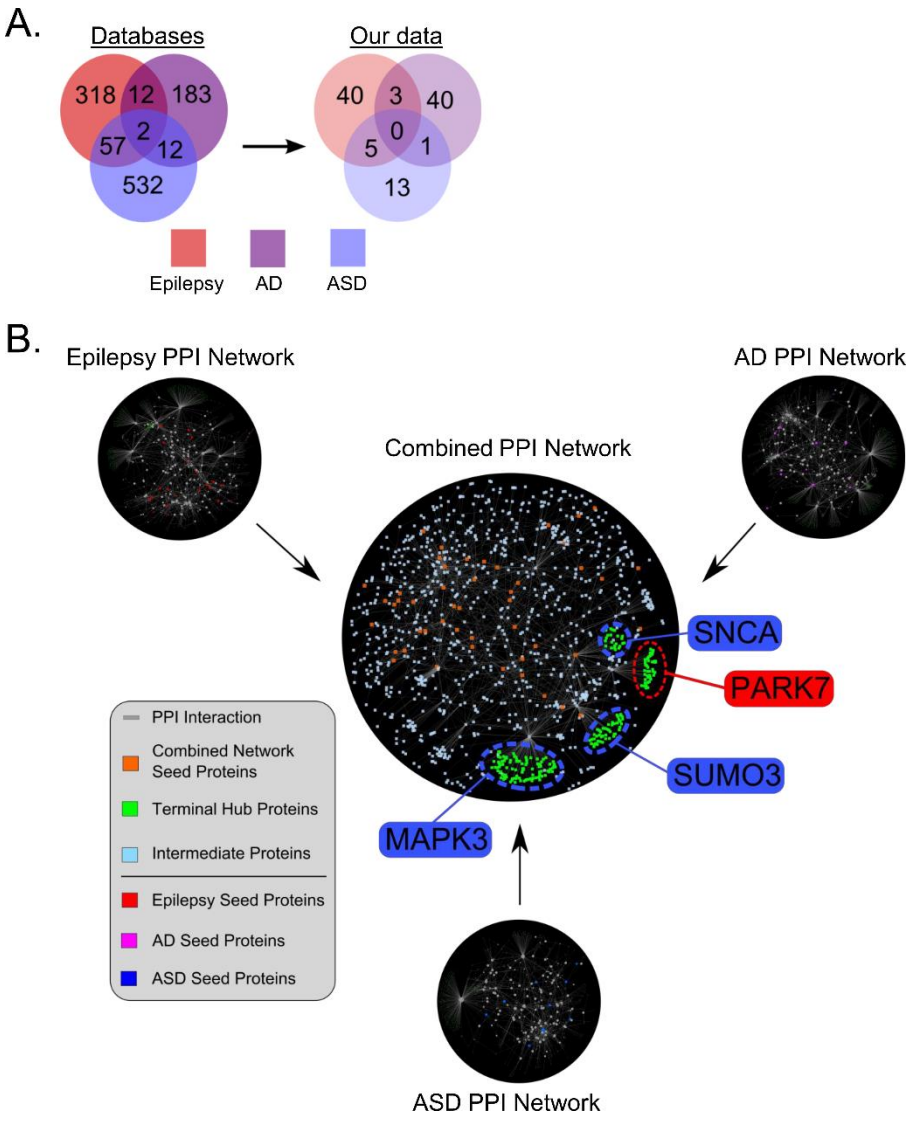


Figure 3.1: Protein-protein interaction (PPI) network analysis determines Parkinson protein 7 (PARK7) as a candidate molecule that may give rise to pathologies common in epilepsy, AD and ASD

A. Venn diagrams depicting candidate genes between epilepsy, Alzheimer's disease (AD), and autism spectrum disorder (ASD). (*Left*, Databases) Venn diagram for genes obtained from disease databases (Table S7A, available online). The genes in the intersection include amyloid precursor protein (APP) and glutamate receptor ionotropic, N-methyl-D-aspartate 2A (Grin2A/GluN2). (*Right*, our data) Subset of left Venn diagram showing only those genes present in our dataset. **B.** Protein-protein interaction networks for epilepsy (*left*), AD (*right*), and ASD (*bottom*) based on

seed proteins up to a distance of two nodes to exact intimate protein interactions (Figure 3.1B, *left*, *right*, *bottom*). Using the degree statistic, we looked for the top nodes that formed the largest number of connections to other nodes (Table 3.1 A1-3), thus allowing us to identify intermediate molecules (herein referred to as hubs) that may link proteins implicated in epilepsy, AD, and ASD. To identify common hubs for all three disorders, we created a network using all 102 proteins identified in these diseases (Fig. 3.1B, *center*). The combined network identified several candidate common hubs (Table 3.1A4, top 10 hubs), with mitogen-activated protein kinase 3 (MAPK3), small ubiquitin-like modifier 3 (SUMO3), α-synuclein (SNCA), and Parkinson protein 7 (PARK7/DJ-1) demonstrating the highest degree statistics in epilepsy, AD, and ASD (Figure 3.1B, *center*). To verify that these common hubs were specific to epilepsy, AD, and ASD, we chose 102 random neuronal proteins from the *R. norvegicus* protein database (NCBI) and created a PPI network under the same parameters. MAPK3, SUMO3, SNCA, and PARK7 hubs were not identified as top candidates by the random PPI network, suggesting that the epilepsy-, AD, and ASD-associated proteins that we have identified from our MS/MS data are uniquely interconnected.

Because these common hubs interact with a high number of proteins, altering the protein abundance of a hub could strongly affect the expression of neighboring proteins, as indicated by their exclusive connections to a cluster of proteins (Figure 3.1B, arrow from each hub pointing to green clusters or terminal hub proteins; Table 3.1). Although two of the top four hubs for the combined PPI network were identified in our MS/MS dataset:

candidate genes. The combined PPI network (*center*) was seeded with epilepsy, AD and ASD candidate proteins found in our data set (Table S7B, available online). Network seed genes are represented by nodes colored in red (epilepsy), magenta (AD), blue (ASD), or orange (combined PPI network). Light blue nodes indicate intermediate PPI interactions between seed genes. Green nodes indicate proteins connected to hub proteins of high degree and/or betweenness centrality (highly trafficked or connected nodes), referred to in the text as “terminal hub proteins”. Proteins around central PPI network indicate proteins connected to the terminal hub proteins including PARK7.

PARK7 and SNCA (Figure 3.1B; Table S10, available online), none of them are currently identified as common to the epilepsy, AD, and ASD databases (Table S7A, available online). Moreover, while SNCA has been identified in AD (Figure 3.1B), PARK7 is absent in all three disorders. Our collective findings highlight the advantage of constructing PPI networks to identify common molecules that may underlie shared pathologic phenotypes of mTORC1-related disorders.

PARK7 levels at the synapse require active mTORC1

Because we identified PARK7 as a novel hub common to epilepsy, AD, and ASD (Figure 3.1B), we performed GO functional annotation to gain insight into how changes in PARK7 expression may alter the expression of terminal hub proteins (Figure 3.2A; Table S10C, available online). Not surprisingly, PARK7 is linked to translation-related processes, suggesting a putative central role in diseases with elevated mTORC1 activity and aberrant translation, like epilepsy, AD, and ASD (Figure 3.2A). Of note, our MS/MS data indicate that mTORC1 activity favors an overall increase in PARK7 throughout the dendrite with the biggest impact at the PSD (Figure 3.2B, heat map). After performing Western blot analysis on cortical PSD fractions isolated from rats that were i.p. injected with carrier or rapamycin for one hour (Figure 3.2C), we found that PARK7 was significantly reduced in the PSD by ~75% when mTORC1 was inhibited (Rapamycin = 0.24 ± 0.10, Control = 1.00 ± 0.06, $p<0.002$), confirming our MS/MS data.

Next, we measured the colocalization of PARK7 with the postsynaptic protein PSD-95 and the presynaptic protein synapsin. Remarkably, rapamycin reduced the normalized PCC of PARK7 and PSD-95 as well as PARK7 and synapsin from ~1.47 to 1.17 and ~1.00 to 0.67, respectively (Figure 3.2D-G). Thus, our MS/MS, Western blot

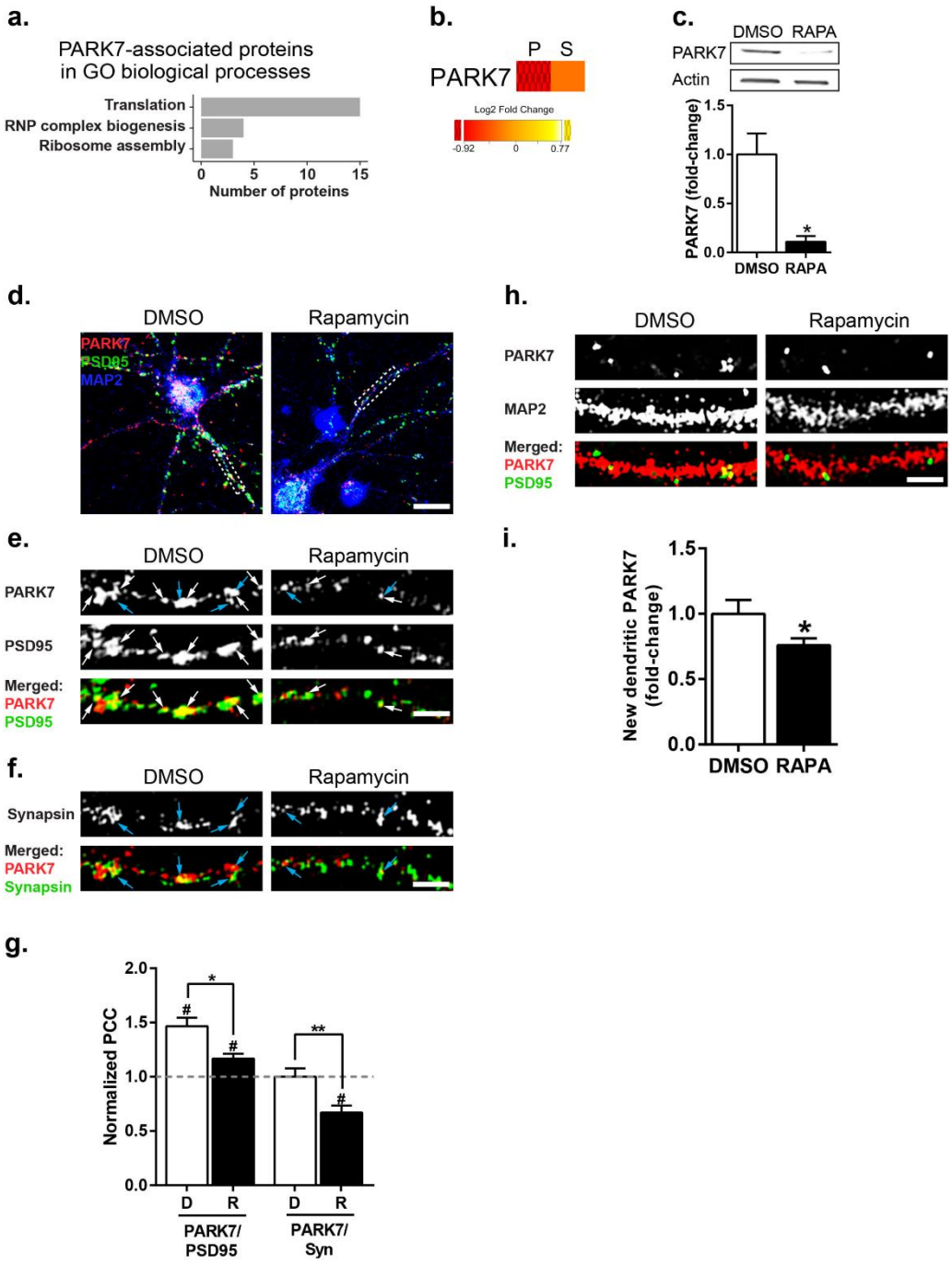


Figure 3.2: mTORC1 activity regulates PARK7 protein expression in PSD and protein synthesis in dendrites

analysis, and colocalization findings show that mTORC1 inhibition reduces PARK7 at the synapse.

If active mTORC1 increases total PARK7 protein through new translation, then rapamycin should reduce synthesis of new PARK7 protein. Surprisingly, using BONCAT-PLA, we found that acute inhibition of mTORC1 reduced synthesis of new PARK7 proteins in dendrites only by ~25% (Figure 3.2H-I; RAPA = 0.76 ± 0.05, Control = 1.00 ± 0.11). While these results demonstrate that mTORC1 positively regulates the translation of Park7 mRNA, the dramatic loss of PARK7 in the PSD also suggest that inhibiting mTORC1 may indirectly destabilize PARK7 protein.

A. Heatmap for PARK7 in different cellular fractions (*top*) and DAVID clustering results for PARK7 terminal hub proteins (*bottom*) (Table S10, available online). Heatmap is shown as $\log_2$ fold change. In PSD, PARK7 is an out-of-range protein with fold-change of 0/Control (*red textured*). **B.** mTORC1 inhibition by rapamycin reduces PARK7 levels in PSD-enriched fraction of rat cortices. Top: representative Western blots. Bottom: densitometric analysis of Western blots. (DMSO = 1.00 ± 0.21; RAPA = 0.11 ± 0.06; n = 3 animals per condition). **C.** Representative images of newly synthesized PARK7 protein when mTORC1 is active (*left, top panel*) or repressed (*right, top*) in or near MAP2 positive dendrites (*middle panels*). Overlay images of new PARK7 proteins and MAP2 (bottom\panels). Scale = 5 μm. See Figure S6, available online. **D.** Rapamycin reduces new PARK7 (DMSO = 1.00 ± 0.11, N = 26 dendrites; RAPA = 0.76 ± 0.05, N = 36 dendrites). Statistics: *, $p < 0.05$; Student's t test. **E.** Representative images of neurons showing differential\ expression of PARK7 (*red*), PSD-95 (*green*), and MAP2 (*blue*). Scale bar 20 μm. **F.** Blown up dendrites indicated by dotted box in **E.** of PARK7 (*top*), PSD-95 (*middle*) and merged image showing colocalization (*bottom*). **G.** Same dendrite as in **F.**, representative staining of synapsin (*top*) colocalized with PARK7 (bottom, merged). Scale bar = 5 μm. PCC is reported directly under merged image for relative comparison between control (*left*) and rapamycin treated neurons (*right*). **H.** PCC averaged over many dendrites for PARK7 and PSD-95 or PARK7 and synapsin. Dotted line indicates average PCC PARK7 and synapsin. Note, PCC of PARK7 with PSD-95 (DMSO = 1.47 ± 0.08, N = 24 dendrites; RAPA = 1.17 ± 0.05, N = 24 dendrites) and synapsin (DMSO = 1.00 ± 0.08, N = 24 dendrites; RAPA = 0.67 ± 0.06, N = 24 dendrites) decreased after rapamycin treatment. Statistics: * and **, significantly different by one-way ANOVA, Tukey's multiple comparison; #, significantly different by single t-test from PCC of PARK7 with synapsin in DMSO. Chicken anti-MAP2 was imaged with 405 nm laser (See Experimental Procedures).

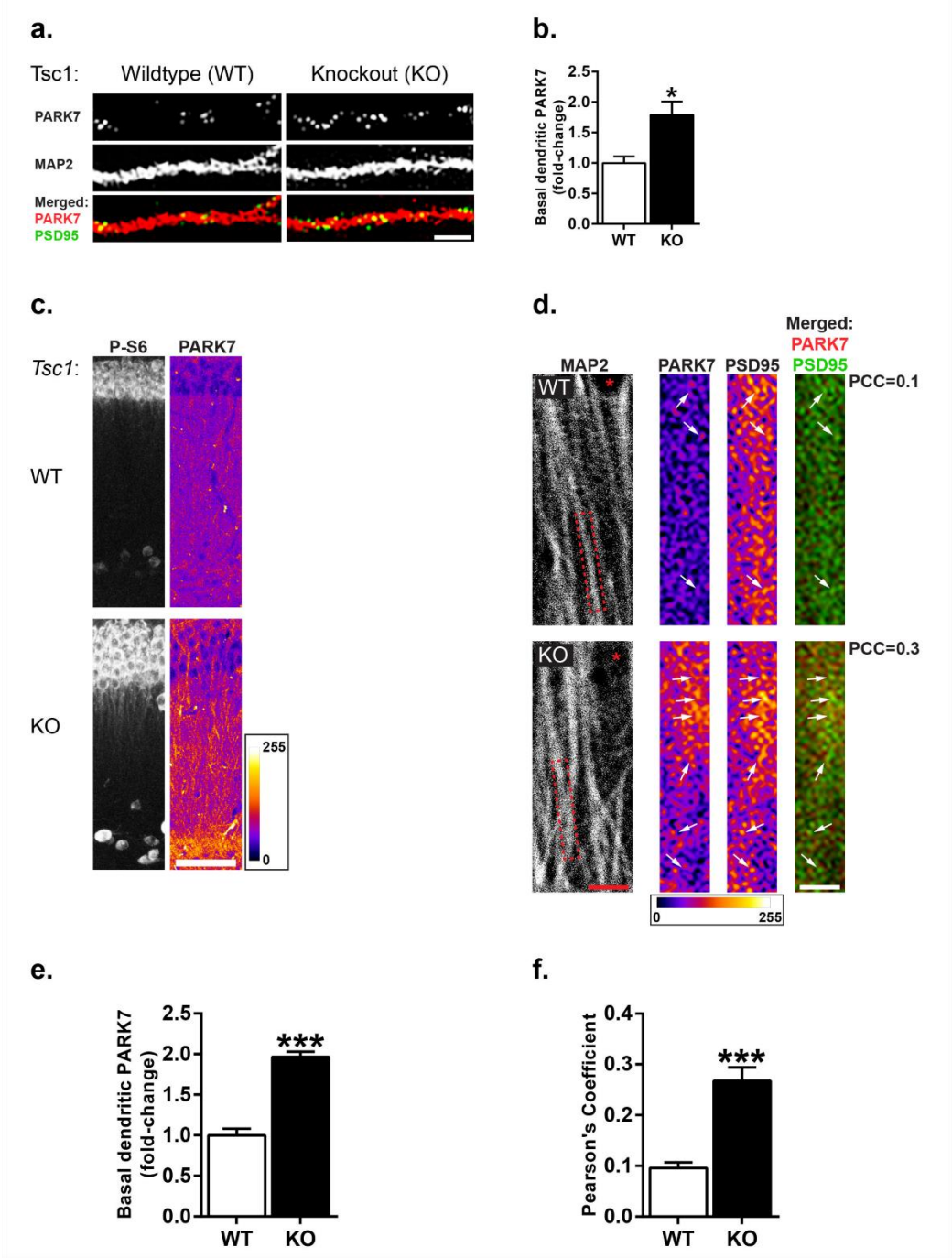


Figure 3.3: PARK7 protein expression is elevated in dendrites and colocalizes with the PSD in *Tsc1* knockout neurons

PARK7 expression is elevated in dendrites of neurons in a mouse model of Tuberous Sclerosis Complex

To test the power of our PPI network to identify new proteins that may be aberrantly expressed in neurological disorders, we used a mouse model of tuberous sclerosis complex (TSC), a disease that exhibits epilepsy and autistic behaviors. In the Tsc1 conditional knockout (cKO) mouse, expression of Cre recombinase virus elevates mTORC1 activity by knocking down the expression of TSC1 protein, an upstream repressor of mTOR. We prepared dissociated hippocampal neurons from Tsc1 cKO mice and infected them at 7-9 DIV with Cre (Tsc1 KO) or GFP (control/WT) for 9-14 days. Using an immunohistochemical approach, we then measured dendritic PARK7 expression at 21 DIV. We found that PARK7 protein in dendrites was markedly increased in neurons infected with Cre (WT = 1.00 ± 0.22, KO = 1.79 ± 0.11; Figure 3.3A-B), suggesting that mTORC1 activation enhances dendritic PARK7 expression.

A. Representative images of basal PARK7 protein (*top panel*) in wildtype (WT) and *Tsc1* knockout (KO) neurons where mTORC1 is overactive in MAP2 positive dendrites (*middle panels*). Overlay images of new PARK7 proteins and dendrites (*bottom panel*). Scale bar = 5 μm. **B.** Quantification of PARK7 protein in or near dendrites. Knockout of *Tsc1* increases basal PARK7 protein (WT = 1.00 ± 0.11, N = 9 dendrites; KO = 1.79 ± 0.22, N = 18 dendrites). Statistics: *, $p < 0.05$; Student's t test. **C.** *Tsc1* cKO mice expressing CRE (KO) or GFP (WT) in the CA1 region of the dorsal hippocampus. Phospho ribosomal S6 (P-S6, *left panel*) staining as a downstream indicator of mTORC1 activity. Slices were costained with PARK7 (*middle*) and MAP2 (*right*). In merged image P-S6 is green, PARK7 is red, and MAP2 is blue. Note enlarged cell bodies in KO, characteristic of overactive mTORC1 signaling. Scale bar = 20 μm. (See Figure S3, available online, for low magnification images). **D.** Merged images of WT and *Tsc1* cKO hippocampal slices stained with PARK7 (*red*), PSD-95 (*green*), and MAP2 (*blue*). Scale bar = 5 μm. Right, dendrite in dotted box in merge image was blown up and colocalization was determined for a single plane (*right*). Note, there is an increase in colocalized pixels with PSD-95 in the KO as noted by the increase in PCC reported next to the merged image. * denotes cell body in WT and KO to demonstrate that blown up images are approximately at an equal distance from cell bodies. Scale bar = 2.5 μm. See Figure S5, available online. **E.** Quantification of basal dendritic PARK7 between WT and KO. Note ~2 fold increase in KO (WT = 1.00 ± 0.08, N = 36 dendrites; KO = 1.97 ± 0.06, N = 36 dendrites). **F.** Quantification of average PCC over many dendrites between PARK7 and PSD-95 (WT = 0.10 ± 0.01, N = 36 dendrites; KO = 0.30 ± 0.03, N = 36 dendrites). PCC increases in KO, consistent with overactive mTORC1 signaling. Statistics: Student's t-test; ***, $p<0.001$. Chicken anti-MAP2 was imaged with 405 nm laser (See Experimental Procedures).

To further extend these findings we stereotaxically injected a recombinant adeno-associated virus (rAAV) for Cre (Tsc1 KO) or EGFP (control) into the hippocampus of adult Tsc1 cKO mice. Immunostaining of hippocampi from these mice showed an increase in the downstream marker of mTORC1 activity, phospho-S6 ribosomal protein, along with the typical increase in cell body size relative to control (Figure 3.3C and Figure S6, available online) (Tavazoie et al., 2005). Of note, the increase in basal levels of dendritic PARK7 in Tsc1 KO mice agreed with our results in cultured neurons (WT = 1.00 ± 0.08, KO = 1.97 ± 0.06; Figure 3.3C, D and E). Moreover, these mice showed more PARK7 colocalized with PSD-95 than control mice (WT = 0.10 ± 0.01, N = 36 dendrites; KO = 0.30 ± 0.03, N = 36 dendrites; Figure 3.3D and F). As predicted by our MS/MS and bioinformatics analyses, our findings demonstrate that in a disease model where mTORC1 is overactive PARK7 expression increases.

Sequence analysis of PSD protein mRNA reveals putative PARK7 binding sites

To test the role of Park7 as a translational hub, we computationally defined a putative PARK7 mRNA binding sequence. Previously, van der Brug et al. combined recovered tags from a PARK7 CLIP experiment to construct a consensus logo that lists the probability of a particular base at each position in the logo (van der Brug et al., 2008) (Figure 3.4A). We took this data and generated all possible combinations of nucleotide sequences of length 6. A length of 6 was chosen since this is a typical length of an RNA-binding motif. We then determined the 6-nucleotide sequence for each position in the 45-nucleotide consensus logo that would have the highest mean probability. Using this method, we derived two motifs of high probability: GNGCNG or CNGCNG. We developed a graphical user interface that allows us to easily scan for the occurrence of thes

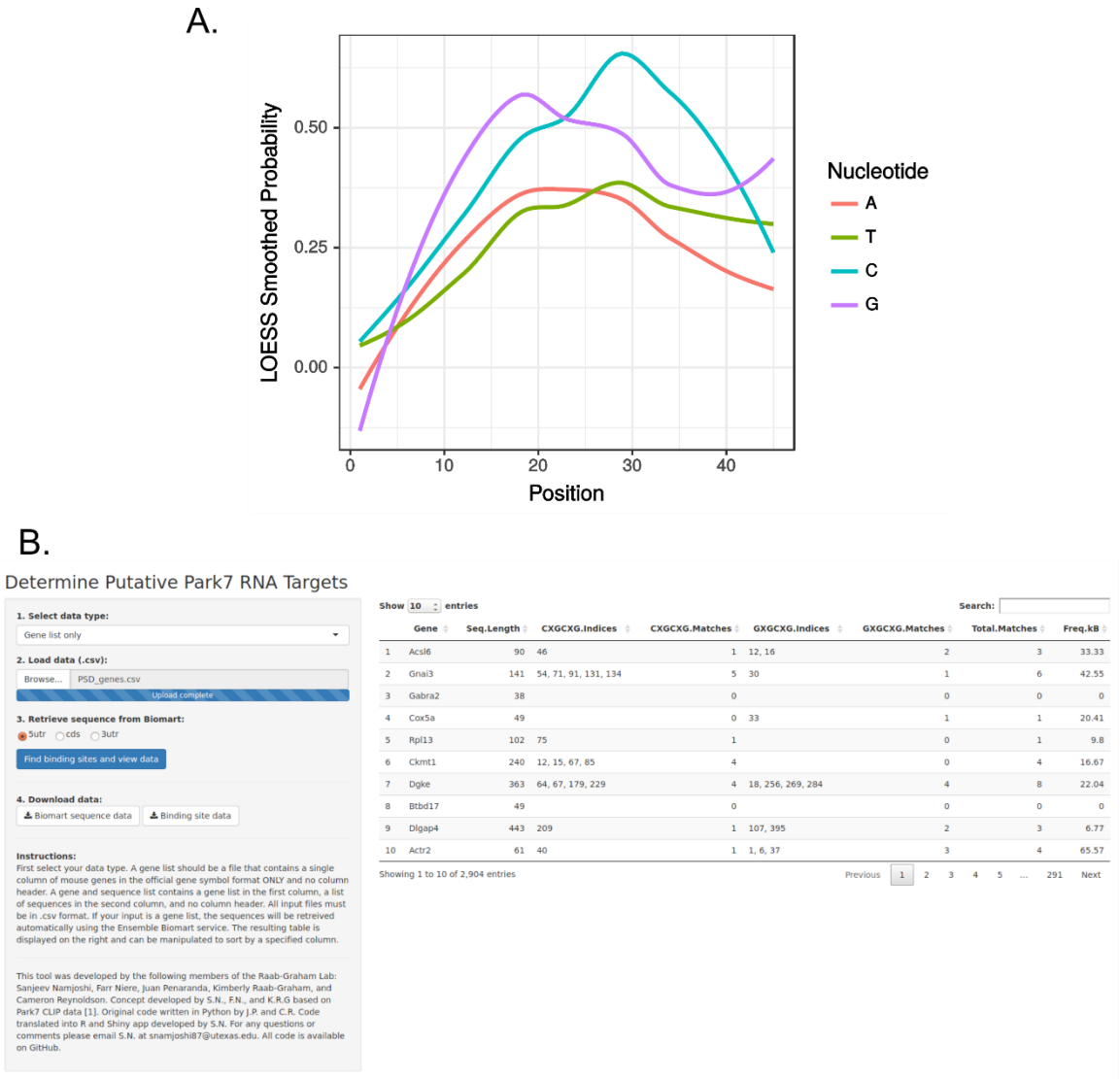


Figure 3.4: Identification of PARK7 binding motifs and calculation of motif frequency

A. Nucleotide probability of PARK7 CLIP tags from van der Brug et al. (van der Brug et al., 2008). Data was LOESS smoothed to interpolate between each position. **B.** Graphical user interface written with the R shiny package that allows for the retrieval of sequences containing the PARK7 RNA-binding motif. Interface accepts gene lists or gene and sequence lists. Uploaded gene lists are sent to Biomart to retrieve the 5'UTR, CDS, or 3'UTR depending on the user's choice. Data output as well as sequence data can be downloaded in .csv format. Interface data table shown on screen can be searched columns may be sorted. The table updates in real time with each of these actions.

sequences within any given input sequence (Figure 3.4B). To test the validity of our tool, we determined the occurrence of the PARK7 motifs in the 5'UTR of the protein kinase Akt since its sequence was recovered by CLIP by van der Brug et al. (van der Brug et al., 2008). We found that the Akt1 5'UTR sequence ranked at the top 82.2% for frequency per kilobase of the PARK7 motifs of the 500 random genes we selected.

Connectivity map analysis reveals potential drugs to modulate expression of mTOR-dependent PARK7 PSD targets

In our previous results, we identified PARK7 as a unique hub connecting together diseases of overactive mTOR (Figure 3.1). Furthermore, our data suggests a role for mTORC1 in regulating the stability of the PARK7 protein (Figure 3.2). We have previously noted that PARK7 expression was highest in the PSD fraction of neuronal cortical tissue by mass spectrometry (Niere et al., 2016). We hypothesized that PARK7 may control expression of PSD proteins that we identified as mTOR targets in our mass spectrometry data. Using PARK7 motifs we scanned mRNA sequences coding for proteins localized to the PSD (Bayes et al., 2012). Out of 2894 5'UTR sequences scanned 21% had putative PARK7 binding sites. Next, we took the list of PSD proteins and filtered them so that they intersected with our mTOR mass spectrometry data. We then obtained the sequences of these mTOR-dependent PARK7 binding targets for the 5'UTR, coding sequence, and 3'UTR and divided them into groups for proteins associated with epilepsy, AD, and ASD (Figure 3.5A). These results show that PARK7 may bind to multiple proteins whose expression is altered by mTOR activity and that PARK7 binds to these mRNA coding for these proteins along different regions of the cDNA. These data provide new avenues of investigation for proteins associated with diseases that may be the result of dysregulated translation in diseases of overactive mTOR.

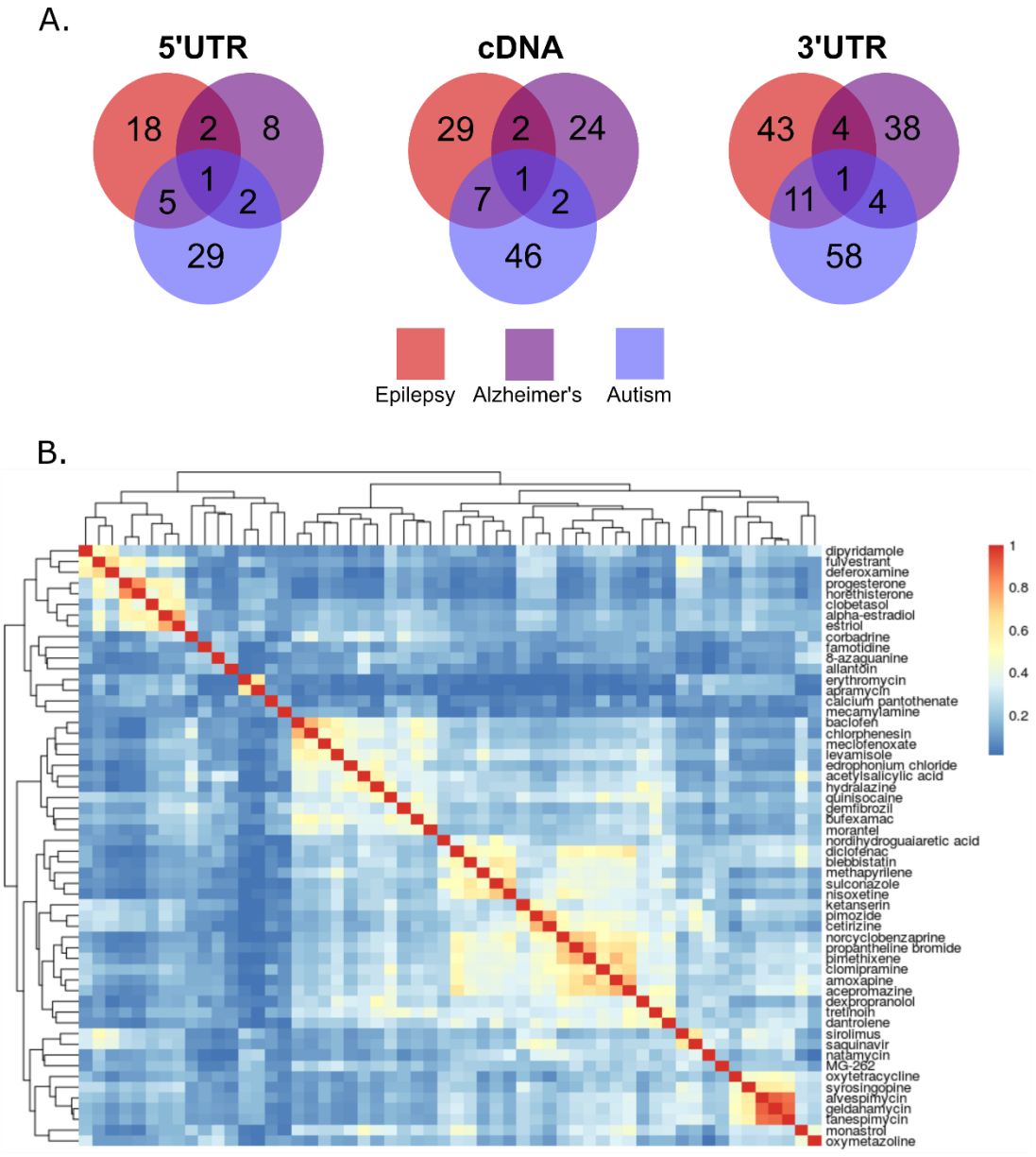


Figure 3.5: PARK7 binding target for disease-associated mTOR MS/MS data and CMap PubChem footprint clustering

A. Number of disease-associated mTOR-regulated proteins with PARK7 binding sites in the 5'UTR, CDS, or 3'UTR. PARK7 binding site data obtained using bioinformatics approach developed in this chapter (see "Experimental Procedures". MS/MS data obtained from (Niere et al., 2016). Sources for disease data shown in Chapter 2 ("Experimental Procedures"). **B.** Hierarchical clustering of PubChem footprints from drugs obtained from CMap. Drugs chosen were above the p-value cutoff of 0.05 from CMap and anti-correlated with the input expression signature.

To further extend this analysis, we asked how pharmacology could be used to reverse the expression changes seen in mTOR-dependent PARK7 proteins. We capitalized on an online tool called ConnectivityMap (CMap) which contains genome-wide microarray data from various cell culture lines after the application of specific drugs (Lamb et al., 2006; Lamb, 2007). We reasoned if we could use the CMap query system to find drugs that produce expression changes anti-correlated with the expression levels we see in our mTOR mass spectrometry data for PARK7 targets, we could produce a potential list of drugs that could reverse changes associated with mTOR-mediated PARK7 dysfunction. We loaded the list of up- or down-regulated mTOR proteins into CMap to produce a list of compounds. Interestingly, rapamycin was the second most highly-ranked compound on the list ($p < 0.0005$) of targets anti-correlated with our input gene signature to CMap. This finding provides evidence to support the validity of our approach. Since the list in CMap is fairly large, we applied a p-value cutoff of 0.05. We obtained the PubChem IDs for the remaining drugs and used the ChemmineR package to determine structural similarities between compounds by applying hierarchical clustering to their PubChem footprint (Figure 3.5B). This segregated the compounds into 13 separate clusters of structural similarity. This approach provides a means for focusing the list to compounds expected to act similarly pharmacologically. Additionally, since many drugs can have side effects, the ability test multiple drugs with slightly different structure provide researchers with the benefit of trying multiple different compounds to minimize these side effects.

DISCUSSION

PARK7 is a putative, dysregulated translational hub in overactive mTOR-related diseases

Individuals with neurological diseases that have dysregulated mTOR activity often suffer spontaneous seizures (Hou and Klann, 2004; Kelleher et al., 2004; Takei et al., 2004; Tsokas et al., 2005; Gong et al., 2006; Antion et al., 2008a; Pei and Hugon, 2008; Swiech et al., 2008; Hoeffer and Klann, 2010; Sharma et al., 2010; Ricciardi et al., 2011; Santini and Klann, 2011; Costa-Mattioli and Monteggia, 2013). We, therefore, looked for common proteins that are associated with epilepsy, AD, and ASD. Surprisingly, we did not detect a single protein shared by available databases for the three disorders that was also in our data set. By creating a PPI network of our disorder-associated data sets, however, we identified PARK7 as a putative protein shared by epilepsy, AD, and ASD. PARK7 is reported to be involved in transcriptional regulation and complexes with processing bodies (P-bodies) suggesting that it may also play a central role in translational regulation as suggested by our PPI network (Ariga et al., 2013; Miller-Fleming et al., 2014). We hypothesize that the PARK7 translational module is aberrantly regulated in diseases with overactive mTORC1. Supporting our hypothesis, we have found increased protein expression in the dendrites of Tsc1 KO neurons, in which mTOR is overactive. Thus, exploring how PARK7 affects the expression of other proteins may represent a previously unexplored avenue for reversing dysregulated protein synthesis in ASD and other mTORC1-related diseases.

Our work, overall, provides evidence that changes in mTORC1 activity induce a rapid and dramatic remodeling of the synaptic proteome by increasing and decreasing syntheses of specific proteins and by altering the local expression of synaptic proteins. However, the precise subcellular targeting mechanisms remain to be determined. We also show that establishing PPI network hubs—common proteins and their associated

molecules—or interactomes sets the stage for identifying proteins, like PARK7, that may account for phenotypic similarities among mTORC1-related diseases. Identification of these common hub proteins can open new avenues to mitigate debilitating phenotypes shared by many neurological disorders.

Remarkably, utilizing protein-protein interaction networks facilitated the novel discovery that PARK7, a protein thus far only implicated in Parkinson's disease, (1) is upregulated by increased mTORC1 activity, (2) resides in the PSD only when mTORC1 is active, (3) is aberrantly expressed in a rodent model of TSC, an mTORC1-related disease that has symptoms of epilepsy and autism, (4) and has multiple putative PSD targets whose expression is known to be dependent on mTOR. Collectively, these data provide the first report of PARK7 association with TSC and its association with multiple diseases characterized by overactive mTOR activity.

SUPPLEMENTAL TABLES

A1. Epilepsy PPI Network Analysis

Protein	Degree	Clustering Coefficient	Terminal Hub Cluster Size
Mapk3	169	0	97
Sumo3	123	0.02	51
Park7	87	0.03	40
Snca	82	0.01	23
Ubc	78	0.01	30
Dlg4	60	0.05	5
Calm1	29	0.2	15
Stx1a	26	0.12	2
Prkca	24	0.14	0
Rad23b	23	0.02	4

A2. Alzheimer's PPI Network Analysis

Protein	Degree	Clustering Coefficient	Terminal Hub Cluster Size
Mapk3	169	0	98
Sumo3	123	0.02	54
Park7	87	0.03	40
Snca	82	0.01	22
Ubc	78	0.01	30
Dlg4	60	0.05	7
Rad23b	47	0.01	19
Slc2a4	30	0	17
Calm1	29	0.2	14
Prkca	24	0.14	0

A3. ASD PPI Network Analysis

Protein	Degree	Clustering Coefficient	Terminal Hub Cluster Size
Mapk3	169	0	98
Ubc	78	0.01	30
Dlg4	60	0.05	5
Snca	35	0.02	5
Sumo3	33	0.06	11
Dlg1	26	0.18	4
Prkca	24	0.14	2
Park7	22	0.14	8
Grip1	22	0.17	0
Magi2	21	0.12	0

A4. Combined PPI Network Analysis

Protein	Degree	Clustering Coefficient	Terminal Hub Cluster Size
Mapk3	169	0	88
Sumo3	123	0.02	46
Park7	87	0.03	39
Snca	82	0.01	22
Ubc	78	0.01	21
Dlg4	60	0.05	5
Rad23b	47	0.01	17
Slc2a4	30	0	15
Calm1	29	0.2	14
Stx1a	26	0.12	2

Table 3.1: Network statistics for PPI networks

A1. – A3. List of top 10 proteins with the highest degree in the PPI network for each of the three diseases. Statistics correspond to the networks shown in Figure 2.6B (*left*, *right*, and *bottom*). Terminal hub cluster size corresponds to the proteins highlighted in green in Figure 2.6B. These proteins have only one neighbor – the protein hub to which they are attached. **(A4.)** Network statistics for combined PPI network (figure 2.6B, *center*).

Chapter 4: mTOR mediates trans-synaptic signaling required for synapse formation and rapid antidepression efficacy requiring Fragile X Mental Retardation Protein[9]

INTRODUCTION

Major Depressive Disorder (MDD) is a chronic illness with a lifetime prevalence of 16.2%, affecting over 30 million US adults (Kessler et al., 2003). The estimated economic burden of MDD has been rising since 2000 when it was estimated at $83.1 billion; it is now estimated to be as high as $210.5 billion with nearly 50% due to workplace costs. Diagnosis of MDD by clinicians typically relies on matching the patient's symptoms to a list of criteria. However, recent work has called these criteria into question on account of the high heterogeneity of symptoms used in diagnosis, low overlap in content between depression scales used by clinicians, and a complex genetic heterogeneity underlying the disease (Lohoff, 2010; Lux and Kendler, 2010; Goldberg, 2011; Hybels et al., 2011; de Vos et al., 2015; Lieblich et al., 2015; Fried, 2017). Research into pharmacological therapies for depression have been under investigation and the most common approach is the usage of SSRIs (Selective Serotonin Reuptake Inhibitors) and other drugs that target molecular pathways of monoamineric neurotransmitters. Yet, recent data from meta-analyses have demonstrated that SSRIs may not be effective in antidepressent therapy – SSRI response rates may be as low as 54% and treatment times may be as long as 12 weeks before a clinical response (Turner et al., 2008; Fournier et al., 2010; Undurraga and Baldessarini, 2012; Greenberg et al., 2015). These findings have motivated the search for alternative pharmacological approaches to treat MDD.

Recently, evidence has suggested that a new class of compounds may be effective in the treatment of MDD. Ketamine, an FDA-approved N-methyl-D-aspartate receptor

[9] For a more detailed introduction see Appendix A

(NMDAR) antagonist has been shown to produce a rapid, long-lasting clinical response in depressed patients thus coining a new class of compounds known as rapid antidepressants (Berman et al., 2000; Zarate et al., 2006). Depression is known to cause loss of dendritic spines and synapse reduction in the hippocampus and PFC (Warner-Schmidt and Duman, 2006; Kang et al., 2012). Ketamine has been shown to reverse this neuronal apathy suggesting its clinical relevance and its potential for treatment of depression (Li et al., 2011). The discovery of rapid antidepressants like ketamine that inhibit NMDAR receptors has prompted research to discover new compounds that may demonstrate similar effectiveness as well as determine their mechanism of action. One such compound is Ro-25-6981 (Ro) which has been shown to have greater specificity than ketamine as it targets the NR2B subunit of NMDARs (Li et al., 2010; Autry et al., 2011; Nosyreva et al., 2013; Workman et al., 2013).

Our recent work has demonstrated that NMDAR antagonism with Ro rapidly activates the serine/threonine protein kinase mTOR through gamma-Aminobutyric acid receptors (GABARs) (Workman et al., 2013; Workman et al., 2015). mTOR is part of the protein complex mTORC1 and has been previously implicated in a number of neurological disorders related to defects in synaptic plasticity (Lipton and Sahin, 2014). Additionally, mTORC1 is a master regulator of translation that controls the expression of hundreds of proteins (Graber et al., 2013b; Niere et al., 2016). It has been shown that mTORC1 signaling is reduced in MDD and that an increase in mTORC1 activity results from treatment with rapid antidepressants (Li et al., 2010; Workman et al., 2013). Furthermore, rapid antidepressant injection leads to an increase in dendritic spine density and this increase is a result of protein synthesis downstream of mTORC1 (Li et al., 2010; Li et al., 2011). Collectively, this work suggests that the reversal of neuronal apathy seen upon

treatment with rapid antidepressants may be the results of an mTOR-mediated remodeling of the synapse.

RNA-binding proteins play an important role at the synapse by coordinating the expression of many mRNA that code for proteins involved in synaptic plasticity (Kiebler and Bassell, 2006; Lenzken et al., 2014; Zhou et al., 2014; Sephton and Yu, 2015). Recently, we have shown that GABAR activation of mTOR though rapid antidepressants leads to a reduction of the RNA-binding protein FMRP (Fragile X Mental Retardation Protein) in dendrites with Ro treatment (Wolfe et al., 2016). Mutations in FMRP result in Fragile X syndrome (FXS), a neurological disorder characterized by aberrant mTOR signaling and dysregulated translation (Sharma et al., 2010; Hoeffer et al., 2012). Additionally, a number of predicted FMRP targets show differential expression changes upon mTOR inhibition (based on data in (Niere et al., 2016)).

To determine the targets controlled by FMRP we devised a high-throughput RNA immunoprecipitation experiment followed by RNASeq (RIP-seq) in mice injected with a rapid antidepressant, Ro-25-6981, and an inhibitor of mTOR, rapamycin. Our work shows that application of the rapid antidepressant is able to shift the RNA population bound to FMRP and furthermore, that this shift is also dependent on mTOR activity. Our work suggests that rapidly-acting antidepressants such as Ro act via presynaptic engagement through mTOR activity that leads to synaptic remodeling. Finally, this work highlights the utility of high-throughput technologies to clarify aspects of neuronal dysfunction and disease.

EXPERIMENTAL PROCEDURES

Injections and experiment design

Wild-type male (C57BL/6, Jackson Laboratory) or *FMR1*-KO (B.6129P2-Fmr1^{tm1Cgr}/J, Jackson Laboratory) mice were injected intraperitoneally (i.p.) with either Ro-25-6981 (10 mg/kg), rapamycin (6 mg/kg), or saline for 45 minutes. Rapamycin cannot be solubilized in saline so we used DMSO at < 1.7% of injection volume based on the animal's mass. This is below the range of DMSO toxicity. Animals were sacrificed after 1 minute under isoflurane. Surgeries were performed to remove the cerebellum, midbrain, and hippocampus.

RNA Immunoprecipitations

Cortices (left hemispheres) from 6-week old WT or *FMR1*-KO mice were harvested and flash frozen on dry ice. Tissue was homogenized and lysed with a cordless pestle motor and disposable pellet mixers (VWR) in polysome lysis buffer (10 mM HEPES pH 7.0, 100 mM KCl, 25 mM EDTA, 5 mM MgCl2, 1mM DTT, 0.5% NP-40) in a 1:1 tissue-buffer ratio. RNaseOUT (Thermo) and protease/phosphatase inhibitors (Halt™ Protease and Phosphatase Cocktail, Pierce Biotechnology) were freshly added to samples. Samples were rotated for 10 minutes at 4°C to induce swelling and then flash frozen on dry ice. Samples were thawed by holding between fingers at RT to lyse and nuclei were pelleted at 3000 x g for 10 minutes. Lysates obtained above were pre-cleared by adding 50 µl of washed magnetic bead slurry (Protein A Dynabeads, Thermo) and rotating for 30 minutes at 4°C. To bind the antibody to the beads, 50 µl of magnetic beads slurry was washed and then resuspended in 8 volumes of NT-2 buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM MgCl2, 1mM DTT, 0.05% NP-40 with RNaseOUT/protease & phosphatase inhibitors

added fresh) + 5% BSA. 10 µg of either FMRP (Abcam, ab17722) or IgG (Santa Cruz Biotechnologies, sc-2027) antibodies were added to the beads and rotated for 10 minutes at RT. Antibody-bound beads were washed 4 times with ice-cold NT-2 buffer. For the immunoprecipitation, 4.5 mg of protein from pre-cleared lysates was added to an RNase-free microcentrifuge tube containing the antibody-bound beads. Input collected at this step for downstream analysis was either 1% of the final pre-cleared lysate volume in the immunoprecipitation reaction (for immunoblotting) or 10% of the final pre-cleared lysate volume (to normalize in qPCR). The antibody-bead-lysate mixture was then diluted at a ratio of 1:5 with NET-2 buffer (20 mM EDTA pH 8.0, and 1 mM DTT in NT-2 buffer; RNaseOUT and protease/phosphatase inhibitors added fresh) and rotated for 1 hour at room temperature. Beads were quickly washed 6 times in ice-cold NT-2 buffer and immediately resuspended in 350 µl TRI Reagent® Solution (Ambion) for 10 minutes at RT. Beads were pelleted and the supernatant was removed and resuspended in 350 µl of absolute ethanol. RNA was extracted by applying ethanol-resuspended samples to spin column from the Direct-zol RNA MiniPrep Kit (Zymogen) according to manufacturer's instructions. Eluted RNA (25 µl) was DNase treated using the TURBO DNA-free kit (Thermo).

cDNA Synthesis and Quantitative Real-Time PCR (qRT-PCR)

15% of the eluate from DNase-treated RNA samples was reverse-transcribed to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad) in a 20 µl volume according to manufacturer's instructions. qRT-PCR was performed in 20 µl reaction volume using the iQ™ SYBR® Green Supermix (Bio-Rad) and primers for $GABA_BR1$, $GABA_BR2$, CaMKIIα, and Cacna2δ2 (GeneCopoeia). qRT-PCR was run with the following protocol:

95°C for 10:00, 40 cycles of 95°C for 0:15 followed by 60°C for 1:00, 95°C for 1:00, and 55°C for 1:00. Relative fold-enrichment was determined by the equation $\Delta\Delta Ct = 2^{-(Ct\ FMRP\ RNA\text{-}IP - Ct\ IgG\ RNA\text{-}IP) - (Ct\ FMRP\ input - Ct\ IgG\ input)}$.

Library preparation

85% of the eluate from DNaseI treated samples were analyzed for quality using the BioAnalyzer (Agilent Genomics). The mean RIN score across all samples was 8.488 (IQR = 8.375 to 8.825) and the mean RNA mass recovered (ng) was 75.31 (IQR = 36.75 to 81.50) (Table 4.2). Recovered RNA was generally lower in the knockout samples but still had a high RIN. Two biological replicates for the Ro knockout samples had a RIN too low for library preparation and sequencing (RIN = 1.0 and 1.1 respectively). Library preparation was performed at the Genome Sequencing and Analysis Facility (GSAF, The University of Texas at Austin). After sample amplification to > 1 ug by PCR, rRNA was depleted using the Ribo-Zero rRNA Removal Kit (Human/Mouse/Rat) (Epicentre) according to the manufacturer's instructions. RNA was then purified using the RNeasy MinElute Cleanup Kit (Qiagen). Library preparation of the RNA was performed with the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (NEB). Single-end sequencing of RNA (SE 1x50) was performed with the HiSeq4000 (Illumina).

Preprocessing of RNA sequence libraries

Quality control assessment of RNAseq libraries

All preprocessing was performed in the R environment (RC, 2014) and terminal commands were sent directly to the shell from R. Code for analysis can be found in Appendix D. RNAseq libraries from all technical replicates were analyzed for quality using

FastQC 0.11.5 (Andrews, 2010) with standard parameter settings. We inspected the libraries for a high Q-score (at least 37) across all base-calls. All libraries showed an overall high Q-score as shown in the "per sequence quality scores" distribution of FastQC. However, multiple libraries showed a small drop in Q-score in either the first six base-calls or the last 2-3 base-calls. The interquartile range for these drops never exceeded a minimum Q-score of 32 demonstrating that overall, the sequencing was of excellent quality. In general, any repetitive sequences were below 1% of all sequences in a given library and usually even lower. A BLAST search from the NCBI website (Boratyn et al., 2013) revealed that they were primarily ribosomal RNA or adaptor sequences. The technical replicates from the Ro knockout library (RO_KO_2) showed the lowest library complexity as well as abnormal GC content and per base sequence content compared to other libraries. Duplicated sequences were shown to be as high as 2%. We analyzed the contents of this library for ribosomal contamination using RNA-SeQC 1.1.8 (DeLuca et al., 2012) but found them to be minimal (< 2% of library sequences). We therefore attribute the unusual sequencing results to poor starting material since these libraries showed the lowest RIN scores and RNA concentration. However, they still contain important sequence information so opted to keep these libraries for our analysis.

Trimming of RNAseq libraries

Since a few libraries showed a lower Q-score toward the beginning and end of the sequences, we applied a trimming tool to the libraries to remove these lower quality base-calls. Trimming would also remove any adaptor contamination since we saw this in a few of the samples. We used the tool Trim Galore 0.4.1 to trim the sequencers and adaptors. Trim Galore utilizes Python 2.7.11 and is wrapped around cutadapt 1.10 (Marcel, 2011).

Trim Galore was run as follows: `trim_galore –fastqc -o trimmed_fasqc *.gz`. Since Trim Galore is pre-loaded with Illumina adaptors we did not need to specify them as parameters. We ran FastQC again to verify that trimming and adaptor removal was carried out successfully and did not alter the quality of the libraries.

Alignment of RNAseq libraries and calculating counts from BAM files

Alignment was performed using Tophat2/Bowtie2 (Langmead and Salzberg, 2012; Kim et al., 2013a) using samtools 0.1.19 (Li et al., 2009a) with the following settings: `tophat2 library-type fr-firststrand -G [GTF] p 4 -o [library] [index] [files]`. The library-type parameter was chosen to be consistent with the library preparation method. Gene model annotations for *M. musculus* (mm9) were downloaded directly from Ensembl (assembly version GRCm38.84, Ensembl 84; Mar 2016) (Cunningham et al., 2015). Pre-built reference indices for Bowtie2 were downloaded from the Illumina website. The alignment data shows that most libraries had a relatively strong alignment percentage (greater than 75%). The only exception is the Ro Knockout replicate 2 (16%). This is not unexpected given the lower library complexity and slightly higher amount of ribosomal RNA contamination. A summary of all library information and alignment data can be found in Table 4.3.

Rsubread 1.20.6 (Liao et al., 2013) was used to calculate counts. Strandedness was maintained in counting using the "`strandSpecific`" parameter. Additionally, we used the "`countMultiMappingReads = TRUE`" parameter option so we could count ambiguously mapped reads. All BAM files for all four technical replicates were combined so that the final output contains counts for all of our libraries.

Post-alignment biological replicate quality control

Count data for each technical replicate were combined together using the `collapseReplicates()` function from DESeq2 (Love et al., 2014). A regularized log transformation was applied to the data before visualization. Data was inspected to see the correlation in count number between different biological replicates (Figure 4.2A), principal components analysis to see the similarity in variance among the replicates (Figure 4.2B), and a Euclidean distance matrix to assess hierarchical clustering of different libraries based on the overall Euclidean distance between each gene (Figure 4.2C).

Filtration/normalization of RNA sequence libraries and differential gene expression analysis

Low count filtration, factor size normalization, and overlap comparison to other FMPR datasets

Combined count data (all technical replicates collapsed) was filtered using HTSFilter 1.14.1 (Rau et al., 2013) with the standard settings. HTSFilter removes all low counts from the full list of raw counts. Annotation was performed using Bioconductor (Huber et al., 2015) with the org.Mm.eg.db (Carlson, 2016c) and AnnotationDbi (Pages et al., 2017) packages. DESeq2 factor size normalization was then applied to the data. Since this will scale across the entire library, the counts in the knockout libraries will be inflated. In order to correct for this, the ratio between extracted RNA concentrations from WT and *Fmr1* KO tissue (BioAnalyzer results) were used to scale back the KO tissue counts after DESeq2 normalization. WT normalized counts were unchanged. All biological replicates were then averaged together across treatments.

To filter the list further, we first sorted the list by the highest counts in the saline treatment and then compared to other known FMRP target data sets. This includes two

RIP-ChIP data sets (Brown et al., 2001; Ascano et al., 2012b), HITS-CLIP data (Darnell et al., 2011), PAR-CLIP data (Ascano et al., 2012b), and APRA-NeuroArray (Miyashiro et al., 2003). Our initial analysis showed poor overlap with the Ascano et al. and Miyahsiro et al. data sets; poor overlap with the Miyasiro dataset was also seen in Suhl et al. (Suhl et al., 2014). The Ascano data also contains a large number of RNA targets (> 5000) and may not be as useful for determining a proper cutoff. These sets were excluded from our analysis, however, consensus datasets showing overlap between the Ascano RIP-ChIP (A-RIP), Ascano PAR-CLIP (A-PAR), Brown (B), and Darnell (D) datasets from Suhl et al. was included in our analysis. Cutoff was determined by sorting the list by highest normalized saline counts and then ranking the percent of matches between the (D), (B) data, and consensus data (B_D_A-RIP, B_D_A-PAR, B_D, and D_A-PAR). By including 4120 genes from our data, we will have a 95% overlap with all genes obtained in the above data (Figure 4.3A). Lastly, we filtered the data by removing mitochondrial and glial cells. These genes were derived from the Brain RNAseq database (Zhang et al., 2014) or MitoMiner (Smith et al., 2012).

Filtering background counts by WT-to-KO fold enrichment

Background was assessed by two methods. 1) We computed the WT-to-KO fold enrichment for all major treatment libraries. Next, we took a list of FMRP non-targets (Darnell et al., 2011) and plotted the WT-to-KO fold enrichment of these non-targets from our data against the WT-to-KO fold enrichment of the targets identified in the other datasets the overlapped in our data (Figure 4.3B). The aim here was to determine a cutoff between the average WT-to-KO fold enrichment for targets versus non-targets, the latter of which were presumed to have a higher background and thus a lower value for the WT-

to-KO fold ratio. We then took the average of the lower quartile of the WT-to-KO fold enrichment across all the different data sets (Figure 4.3B, dotted red line) and used this as an estimation of the cutoff for background. 2) We used univariate K-means with the Ckmeans.1d.dp package (Wang and Song, 2011) was used to quantify the distribution of signal and background counts into two clusters (Figure 4.3C, D). Gaussian mixture modeling (GMM), a generalization of K-means, has been used previously to characterize signal-to-noise for RIP-SEQ/RIP-ChIP data (Erhard et al., 2013; Lu et al., 2014), however our data did not converge using GMM. We then averaged the cutoffs from the two methods used above for WT-to-KO fold enrichment ratio cutoff of ~1.112788. We applied this value across all treatments to filter any data that did not meet this cutoff criterion.

Differential gene expression analysis

Differential gene expression (DGE) analysis was performed using Limma/voom (Law et al., 2014; Ritchie et al., 2015). Raw counts were imported into R and then normalized via TMM/CPM. The contrast matrix was set up to compare fold changes for either Saline/Ro (SAL_WT – RO_WT) or Ro+Rapa/Ro (RO_RAP_WT – RO_WT). After running the Limma/voom pipeline, samples were thresholded with an FDR cutoff of 0.1. These values were then visualized using an MA plot (Figure 4.4A) and a volcano plot (Figure 4.4B).

Motif analysis of FMRP targets

For motif analysis, we did not use the differentially expressed genes selected by Limma/voom because motif frequency/number was not significantly higher than that of the non-targets. All sequences were obtained from BioMart (Smedley et al., 2015) using the

biomaRt (Durinck et al., 2005; Durinck et al., 2009) package from Bioconductor. Sequences were obtained from the coding region because the majority of FMRP binding motifs appear to cluster in these part of the gene (Anderson et al., 2016). Since some genes may return multiple sequences, we only used sequences of max length for our calculations. Motifs were counted using the `count()` function from the seqinr (Charif and Lobry, 2007) package. The QFM sequence was counted using the `matchPDict()` function from the biostrings (Pages et al., 2016) package using the following approach: the "WGGA" motif is composed of sequences of either "AGGA" or "TGGA". Both sequences are searched for within a given sequence string. Then, the start position of either motif within the sequence is ordered and the difference between the start and end positions is determined by taking the difference across each row. Finally, the QFM motif is determined by subsetting this dataframe for motifs that are only separated by a span of less than 6 bases (based on the difference between the start and end positions). Rho is computed using the `rho()` function from seqinr. To compute multiple rho, which is necessary for motifs for which a position can take on more than one nucleotide value (e.g. ACUK), a function called `multipleRho()` is used which computes rho manually according to the following formula:

$$\frac{[frequency(motif1) + frequency(motif2)]}{[expected(motif1) + expected(motif2)]}$$

Motif statistics were calculated using base R functions except for the effect size which uses the `cohensD()` function from the lsr package (Navarro, 2015). For all motif statistics (freq/kB, t-test, Wilcoxon-Mann-Whitney test, effect size) except rho, we followed the general presentation of data from Suhl et al. (Suhl et al., 2014).

Clustering analysis of FMRP targets

GO clustering of unique biological processes

For clustering analysis, except where specified, we did not use the differentially expressed genes selected by Limma/voom because motif frequency/number was not significantly higher than that of the non-targets. Database for Annotation, Visualization and Integrated Discovery (DAVID 6.8) (Huang da et al., 2009a; 2009b) was queried for gene ontology annotation using the RDAVIDWebService package (Fresno and Fernandez, 2013). To cluster GO terms with high similarity we used an approach similar that of the Enrichment Map plugin (Merico et al., 2010; Merico et al., 2011) for Cytoscape 3.0 (Shannon et al., 2003). First, we obtained DAVID output from their functional annotation chart service from the "GOTERM_BP_ALL" ontology. The up-regulated and down-regulated genes were analyzed separately by DAVID. Then, we filtered these GO terms using a p-value cutoff of 0.001 and an FDR cutoff of 0.05. We then took the relative complement of either the up-regulated or down-regulated GO terms. We then use a clustering approach to group GO terms of high similarity (based on the number of overlapping genes) from each relative complement together in a dendrogram (Figure 4.10). To determine the degree of overlap between genes in any two ontologies we used the average of the Jaccard cutoff (JC) and the overlap coefficient (OC).

$$JC = \frac{|A \cap B|}{|A \cup B|}$$

$$OC = \frac{|A \cap B|}{Min(|A|, |B|}$$

A higher overlap score indicates a greater number of genes in common. We next created a matrix of the average JC and OC for every GO-term against all other GO-terms. The matrix

was filtered for a similarity cutoff of > 0.5. This matrix was then clustered to create a dendrogram using the factoextra package (Kassambara and Mundt, 2016). Each branch of the dendrogram indicates a degree of overlap between the different GO terms. Each clustered branch was then manually annotated for a general biological category that described the GO term. To determine optimal cluster number, we tried to minimize the within-sum-of-squares (WSS) (Figure 4.11).

Univariate/bivariate K-means clustering and clustered heatmaps

Univariate clustering was performed as described previously using WSS to determine optimal cluster number (Figure 4.11). GO clustering from each cluster used the RDAVIDWebService package. For bivariate K-means clustering we used the `Kmeans()` function from the amap package (Lucas, 2014). To create the fold-change correlation plot, the all genes in the interquartile range were colored grey to indicate "background" or no change. Trannsynaptic proteins were obtained from the gene ontology annotations (see "Connectivity map analysis of FMRP targets" section below). Clustered heatmaps of the limma-filtered data was created using the pheatmap package (Kolde, 2015). All data outside of the range of -2 to 2 $\log_2$ fold-change were removed before plotting. Only data with an adjusted p-value less than 0.1 were included. For the heatmap of the normalized count data, the data was $\log_2$ transformed before plotting with pheatmap.

Depression GO terms chart data source

All of the genes from the depression GO terms chart were obtained through an online literature search for genes that has been associated with depression (Fatemi et al., 2001; Tadic et al., 2007; Dlugos et al., 2009; Unschuld et al., 2009; Lee et al., 2010b; Luo

et al., 2010; Rietschel et al., 2010; Aragam et al., 2011; Galeotti and Ghelardini, 2011; Shyn et al., 2011; Dekker et al., 2012; Sakaida et al., 2013; Ray et al., 2014; Gatt et al., 2015; Gray et al., 2015; Ogawa and Kunugi, 2015; Skoog et al., 2015; Wagner et al., 2015; Yoshimasu et al., 2015; De Vry et al., 2016; Liu et al., 2016; Stachowicz et al., 2016).

Usage of other R packages for analysis

In addition to the packages already mentioned, we utilized a number of other packages for our analysis. For plotting: ggplot2 (Wickham, 2009), ggthemes (Arnold, 2016), RColorBrewer (Neuwirth, 2014), and gridExtra (Auguie, 2016). For annotation: mgu74a.db (Carlson, 2016a) and org.Hs.eg.db (Carlson, 2016b). For functions used in DGE analysis: edgeR (Robinson et al., 2010). For data import: xlsx (Dragulescu, 2014) and R.utils (Bengtsson, 2016). For data transformation and processing: reshape2 (Wickham, 2007), magrittr (Bache and Wickham, 2014), dplyr (Wickham and Francois, 2016), and gdata (Warnes et al., 2015).

RESULTS

Gene Ontology (GO) enrichment reveals high heterogeneity of biological processes for candidate depression genes

Symptoms of patients with MDD show a high level of heterogeneity between diagnosis criteria. To see if this heterogeneity was represented on the molecular scale we performed GO biological process analysis on a list of candidate depression genes derived from a literature search and used the Database for Annotation, Visualization and Integrated Discovery (DAVID 6.8) to cluster these genes. The resulting list was then grouped by unsupervised hierarchical clustering into simplified categories based on the similarity of overlap between the genes shared in the biological processes returned by DAVID (see "Experimental Procedures"). Our analysis revealed over 30 major categories with little overlap in functional properties such as hormone secretion, circadian rhythm regulation, cytokine production, and regulation of cell motility consistent with previous findings regarding the heterogeneity of MDD symptoms (Figure 4.8A). Furthermore, the same GO analysis of genes from two other datasets that isolated mRNA bound to FMRP (Brown et al., 2001; Darnell et al., 2011) shows little overlap with the MDD GO biological process chart (Figure 4.8B). This suggests that if FMRP targets related to depression have been identified previously, they did not significantly cluster into GO categories that overlapped with MDD genes. Since our previous work has shown that dendritic FMRP levels drop significantly after administration of rapid antidepressants \\Wolfe, 2016 #332//, we reasoned that immunoprecipitating FMRP target mRNAs may reveal a population that could be utilized to develop more targeted treatments for MDD.

RNA-immunoprecipitation of FMRP mRNA targets in the mouse cortex

The large diversity in unrelated biological processes provides a great challenge for the proper design of pharmaceuticals to treat MDD. We reasoned that we could optimize strategies for antidepressant drug discovery by characterizing the molecular changes that occur at the synapse upon treatment with rapid antidepressants. We sought to identify the mRNA targets that show differential binding to FMRP upon rapid antidepressant treatment. We performed RIP-seq in mice injected with either Ro and the mTOR inhibitor rapamycin, Ro alone, or vehicle for three biological replicates (Figure 4.1). RIP-seq was performed in wild-type (WT) mice as well as the Fmr1-knockout (Fmr1-KO) mice. Our RIP-seq approach used a number of modifications that have been shown previously to significantly reduce background mRNA binding (Keene et al., 2006; Jain et al., 2011). We sequenced four technical replicates for each RNAseq library resulting in WT libraries ranging from 16-27 million reads (mean = 22.1 million reads, s.d. = 4.3 million reads) (Table 4.3). Reads obtained in KO libraries were highly variable ranging from 8-35 million reads (mean = 19.6 million reads, s.d. = 10.3 million reads). However, reads from KO libraries mapped lower than those of their WT counterparts (mean KO library mapping rate = 57.64% ± 21.11% (s.d.); mean WT library mapping rate = 68.94% ± 18.59% (s.d)) (Table 4.3). In general, our libraries were of high quality and showed high replicability.

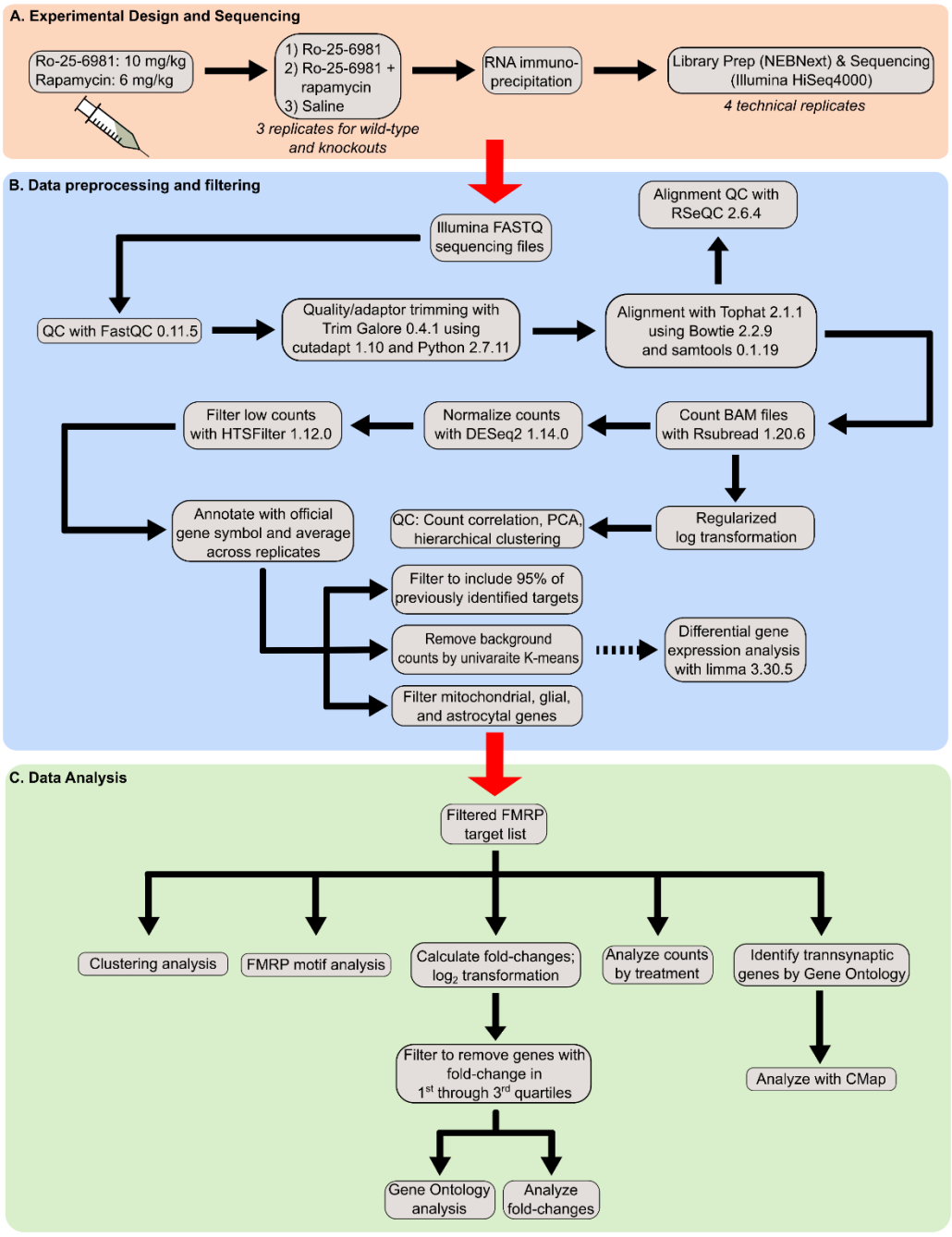


Figure 4.1: Workflow for RIP-seq experiment and analysis

We visualized the similarity between replicates after applying a regularized log transformation to the raw counts (Love et al., 2014). Biological replicates between libraries were highly correlated which demonstrates high precision in RNA sample extraction and preparation (Figure 4.2A). Next, we created a Euclidean distance matrix to see the similarity in counts across all observations for a given biological replicate (Figure 4.2B). Replicates within a given treatment generally clustered into the same groups, the greatest exception being the second and third replicates of the Saline treatment library (SAL_WT2; SAL_WT3). However, all saline replicates are still within the same parent branch. Next, we examined the sample similarity through principal components analysis (PCA) (Figure 4.2C). Biological replicates within the same library generally group together. The most notable exception is the 2nd Saline KO library (SAL_KO2). Taken together, the high overlap of similar groups by PCA and clustering as well as the high correlation between replicates within a treatment group demonstrate the reliability of our data through different repetitions of the experiment.

To verify the validity of our sequencing results, we first took a list of FMRP mRNA targets and non-targets from a previous publication from the Darnell lab and compared the counts for these targets obtained from the WT or Fmr1-KO mice across treatments (Figure 4.9A) (Darnell et al., 2011). We found the normalized counts for the targets were increased by 333.90% (Saline WT), 369.29% (Ro WT), and 365.35% (Ro + rapamycin WT) over non-targets in each treatment. This supports the validity of our obtained libraries. Furthermore, we inspected the result of the normalization to ensure that the ratio between WT and KO counts for a given treatment had been maintained. Since this was performed prior to low-count filtering, the counts are still highly zero-inflated (Figure 4.9B). However, there is a clearly a higher number of zero counts in the WT compared to KO counts within specific treatment. Next, we performed qRT-PCR on our

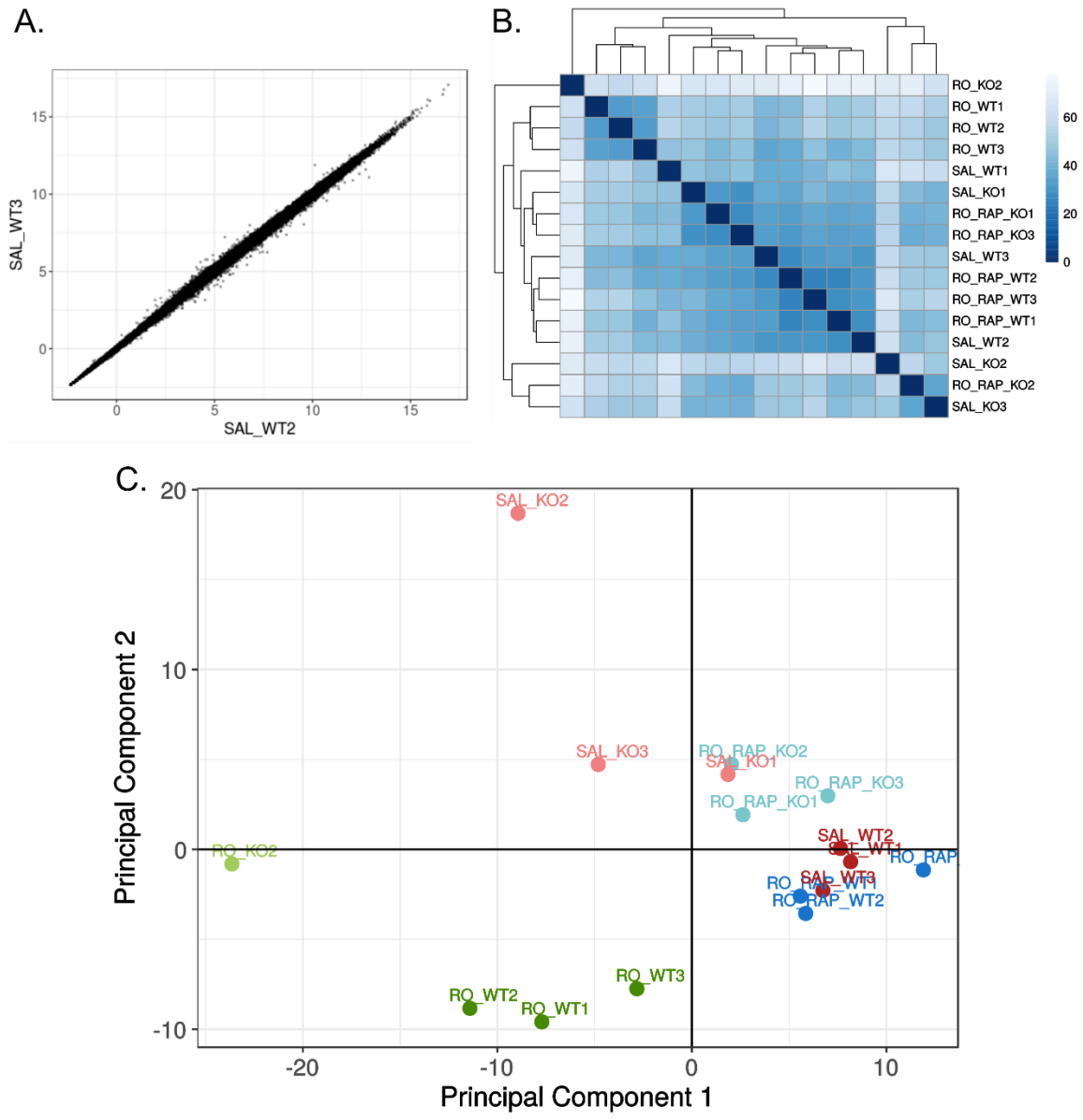


Figure 4.2: Quality control of count data from RIP-seq libraries

All counts analyzed were transformed with a regularized log transformation of raw count data for each library. **A.** Correlation of a representative RIP-seq libraries indicates high correlation between targets. **B.** Euclidean distance matrix shows high similarity between RIP-seq libraries using the complete linkage method for clustering. **C.** Principal component analysis of observations in the RIP-seq libraries. Color shades indicate WT (dark shade) or KO (light shade) libraries.

immunoprecipitated RNA using primers for $GABA_BR1$ and $GABA_BR2$, two FMRP targets we have previously characterized, under Ro or Saline treatments (Figure 4.9C). This data confirms the presence of $GABA_BR1$/$GABA_BR2$ mRNA in our samples and shows that their levels decrease with Ro treatment.

Filtration of RNA-IP data results in a high-quality dataset enriched for multiple neuronal GO clusters and differentially expressed FMRP targets

Candidate RNA targets of FMRP have been obtained previously by numerous different methodologies including RIP-ChIP, HITS-CLIP, and PAR-CLIP (Brown et al., 2001; Miyashiro et al., 2003; Darnell et al., 2011; Ascano et al., 2012b). We leveraged the results from these studies to filter our data to obtain a list of FMRP targets based on a consensus overlap between these data sets following the work by Suhl et al. (Suhl et al., 2014). Here, we attempted to capture 95% of the FMRP mRNA targets present in these overlap data sets for our candidate targets with the highest counts (Figure 4.3A). The Ascano et al. and Miyashiro et al. datasets by themselves were not included in this analysis due to insufficient overlap with our data. This finding was similar to Suhl et al. who excluded the Miyashiro et al. dataset as well (Suhl et al., 2014). After filtering further to remove any mitochondrial or glial contaminant mRNA from our data based on an online database (Cahoy et al., 2008) we further filtered by establishing a count cutoff for target v. non-targets by determining the count enrichment of FMRP WT mRNA relative to Fmr1-KO mRNA (Figure 4.3B). Using the FMRP target mRNA from the consensus list from (Suhl et al., 2014) and non-target mRNA from (Darnell et al., 2011) we searched for genes in common with our candidate FMRP target list and determined the distribution of fold changes in the saline WT condition compared to saline KO condition. Taking the bottom quartile across all target groups yielded a WT/KO enrichment cutoff of ~1.11. We used

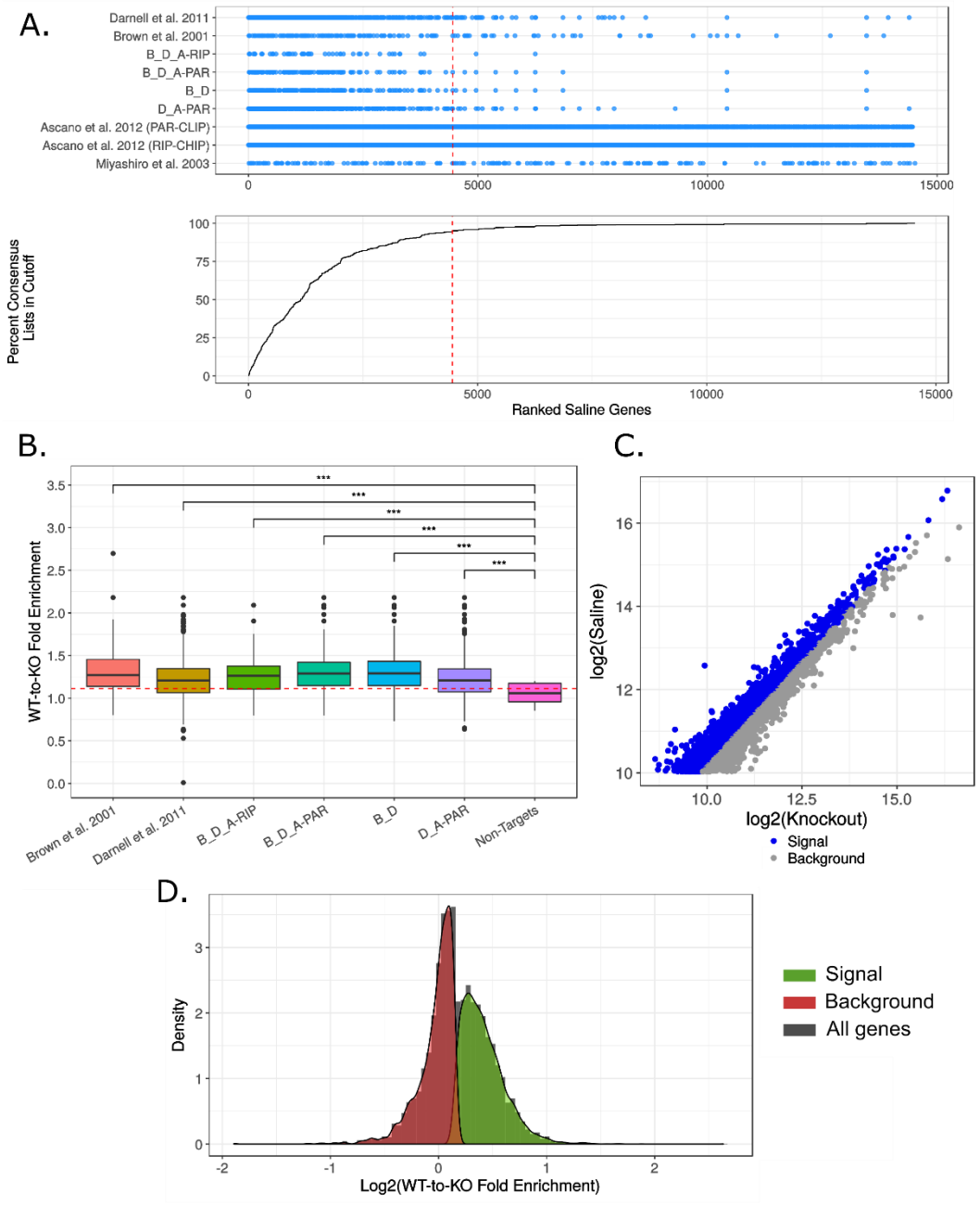


Figure 4.3: Filtration of dataset using FMRP consensus data set targets and background from RIP-seq KO libraries

A. *Top:* Distribution of FMRP targets from other datasets across our data. Each position along the X-axis represents a gene found in the Saline RIP-seq library. The gene with the highest count has a position near the origin while the gene with the lowest count has a position at the other end of the axis. The Y-axis indicates a data from the Ascano PAR-CLIP data (A-PAR), Ascano RIP-seq data

this metric to further filter our list of FMRP mRNA targets. To provide a more rigorous, unbiased filter for WT/KO enrichment in our data, we used K-means clustering to find a sub-population for background counts within our data. This approach is similar to Gaussian mixture modeling which has been previously used to analysis RIP-ChIP and RIP-Seq data (Erhard et al., 2013; Lu et al., 2014). This analysis revealed two separate populations among the $\log_2$ transformed WT/KO enrichment data – one population consisting of background ($\mu = 0.985$, $\sigma = 0.109$) and a second population consisting of the FMRP mRNA targets ($\mu = 1.352$, $\sigma = 0.23$) (Figure 4.3C,D). The maximum WT/KO enrichment ratio found in the background population is less than 1 thus providing a similar cutoff to that obtained by comparing to targets and non-targets as described above. Using the data from these distributions, we visualized the separation of enrichment between target and background through a scatterplot (Figure 4.3C). Altogether, this approach allowed us to filter out any mRNA from our data that was not enriched relative to the counts obtained in KO tissue.

(A-RIP), Darnell data (D), Brown data (B), or Miyashiro data. See text for data citations. Consensus data sets indicate overlap between the above data (e.g. B_D_A-PAR indicates the intersection between the Brown, Darnell, and Ascano PAR-CLIP data). These overlap sets were obtained from the supplemental material in Suhl et al. (Suhl et al., 2014). Dotted red line indicates cutoff that was used in our data. Everything to the left of the line is considered a putative FMRP target and everything to the right is considered a non-target. *Bottom:* Line graph indicating the percent of consensus FMRP targets (excluding the Ascano and Miyashiro data) that overlap with our data for each position in the corresponding graph above. Dotted red line indicates the same as in **A.** For all putative FMRP targets (left of dotted red line), 95% of consensus FMRP targets are captured. **B.** WT-to-KO fold enrichment for genes in our data that overlap with a data set indicated on the X-axis. Non-targets were obtained from (Darnell et al., 2011). Dotted red line indicates the average bottom quartile for all FMRP target data sets. According to this cutoff, everything above the line is a putative target and everything below is considered background binding. Statistics: ***, $p < 0.001$; Student's t-test. **C.** Scatterplot showing the relationship between the WT and KO counts in the $\log_2$ saline RIP-seq library. Cutoff between "signal" or "background" determined by univariate K-means. **D.** Univariate K-means for $\log_2$ saline counts. The two clusters generated (signal, green; background, red) are overlaid on a histogram of the full saline distribution (grey). The max value from the background cluster was chosen as the cutoff point for the distinction between signal and background.

After obtaining our filtered data set above, we performed differential expression analysis (DGE) using limma+voom (Law et al., 2014; Ritchie et al., 2015). Using an FDR cutoff of 0.1 we obtained 242 genes across either Saline/Ro or Ro+Rapamycin/Ro treatments (Figure 4.4A,B). We predict that saline/Ro fold changes will identify FMRP targets translated with Ro induction ($\log_2$ fold-change > 1) or FMRP targets not translated with Ro induction ($\log_2$ fold-change < 1). Furthermore, we predict that the Ro+Rapamycin/Ro fold changes will identify FMRP targets that are translated when Ro turns mTOR on ($\log_2$ fold-change > 1) or FMRP targets that are repressed when mTOR is on ($\log_2$ fold-change < 1). Next, we performed hierarchical clustering to observe mRNA that changed together in both fold-change groups (either Saline/Ro or Ro+Rapamycin/Ro). (Figure 4.4C). In general, the majority of mRNA were not differentially expressed, a relationship that we explore in greater detail below (Figure 4.6, 4.7). The genes showing the greatest differential expression when the fold-change groups were compared this way were primarily mRNA associated with the small and large subunits of the ribosome. We also clustered all libraries prior to averaging by performing a regularized log transformation of the raw counts (Figure 4.4D). GO clustering for biological process revealed six enriched clusters which included processes associated with trans-synaptic signaling and nervous system development. Notably, we observed differential binding between the different treatments indicating that FMRP sequesters a separate set of targets upon administration of Ro (Figure 4.4E).

Attempts to characterize FMRP binding motifs have identified multiple common patterns among targets including the G-quadraplex motif, U-rich sequences, and other common motif patterns (Schaeffer et al., 2001; Dolzhanskaya et al., 2003; Menon and

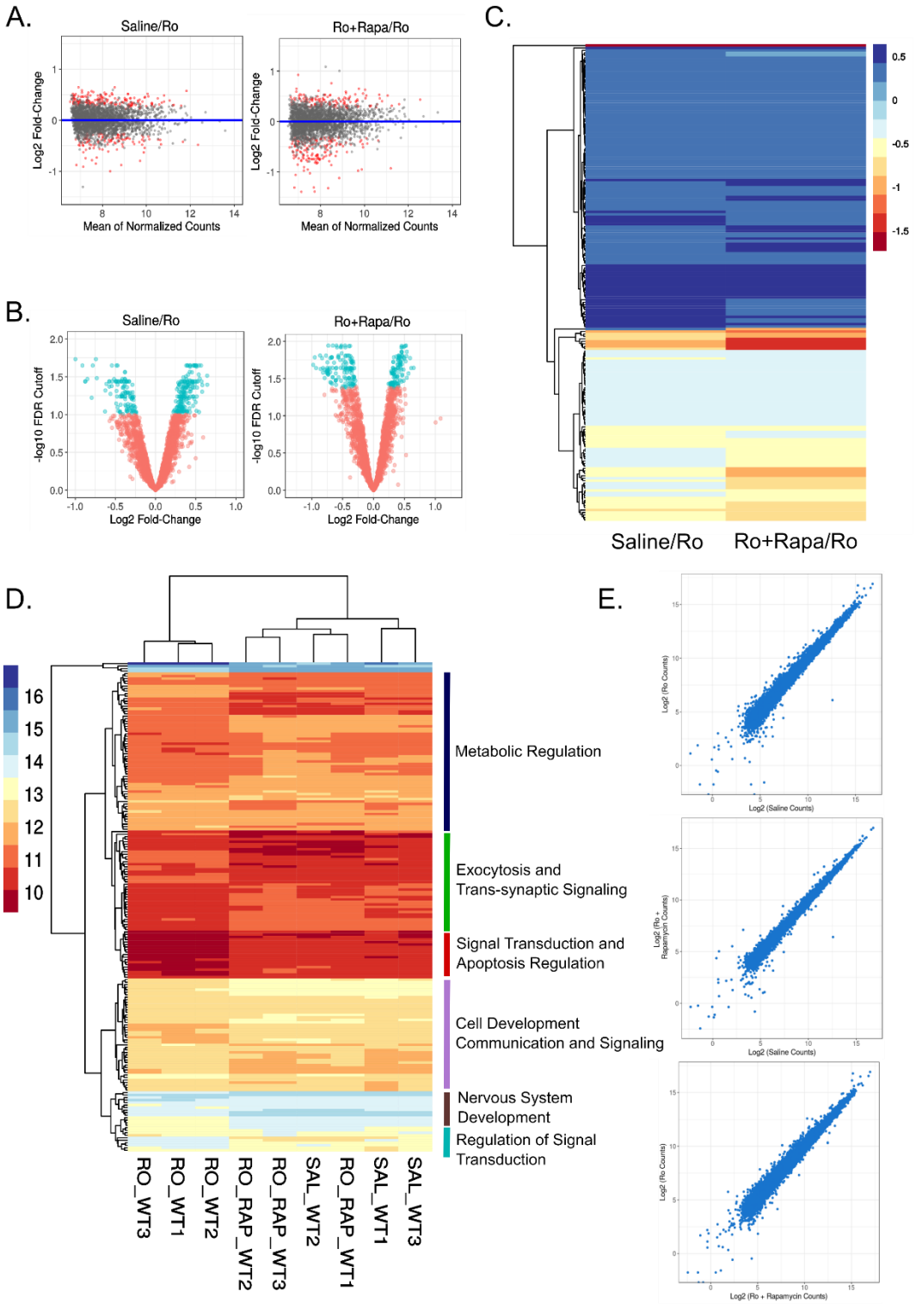


Figure 4.4: Differential gene expression analysis of RIP-seq libraries

Mihailescu, 2007; Menon et al., 2008; Darnell et al., 2011; Phan et al., 2011; Ascano et al., 2012b; Ray et al., 2013; Suhl et al., 2014; Chen and Joseph, 2015; Anderson et al., 2016). Although there is no consensus on which patterns are common to all FMRP targets, Suhl et al. have attempted to bioinformatically determine the occurrence of all these motifs within known FMRP datasets. Our approach was to partially replicate their overall analysis strategy with our targets to see if we obtained consistent results. We characterized the enrichment of putative FMRP target motifs across our filtered dataset compared to genes that had been filtered out of our dataset in previous processing steps described above (Figure 4.5A; Table 4.1). We also calculated the occurrence of the motif within sequences that were filtered out of the list ("non-targets"). We consider a motif to be enriched in our FMRP targets (n = 2696) if the frequency of occurrence per kilobase (kB) was significantly higher than the occurrence in non-targets (n = 11222). Sequences were used specifically from the coding region because the majority of FMRP motifs are found here (Anderson et al., 2016). Additionally, we also calculated ρ which tells us if the occurrence of a DNA string is over- or under-represented according to what we would expect by chance (see "Experimental Procedures"). For example, ρ = 10 would indicate the occurrence of the sequence is 10 times more likely that we would expect by chance.

Data for motif analysis is summarized in Table 4.1. The mean frequency/kB of the ACUK motif in the FMRP targets was not significantly enriched compared to the non-

A. MA plot showing distribution of counts by fold-change for each of the $\log_2$ transformed fold-change conditions under consideration. Significant genes are shown in red and non-significant genes are shown in grey according to an FDR cutoff of 0.1. **B.** Volcano plots for the fold-change conditions under consideration. Significant genes are shown in blue and non-significant genes are shown in red. **C.** Clustered dendrogram for fold-change conditions using the average linkage method and Euclidean distance. **D.** Clustered dendrogram for each RIP-seq library after regularized log transformation. Scale represents rlog transformed raw counts. Clustering performed using the average linkage method and Euclidean distance. Colored boxed on the right side correspond to GO biological process clusters named below the chart. GO clustering performed with DAVID. **E.** Scatterplot visualizing the distribution of counts between different RIP-seq libraries after $\log_2$ transformation. Scale represents $\log_2$ transformed normalized counts.

Motif	Freq/kB (Targets)	Freq/kB (Non-Targets)	ρ (Targets)	ρ (Non-Targets)	*t*-test	Wilcoxon-Mann-Whitney Test	Cohen's *d*
ACUK	8.036	8.483	0.455	0.477	$p < 0.0001$	$p < 0.0001$	0.144
WGGA	16.531	16.666	0.200	0.201	$p = 0.222$	$p = 0.239$	0.026
GAC	16.787	16.066	0.955	0.936	$p < 0.0001$	$p < 0.0001$	0.162
GACR	7.905	7.444	0.222	0.214	$p < 0.0001$	$p < 0.0001$	0.152
GACARG	0.875	0.773	0.024	0.022	$p < 0.0001$	$p < 0.0001$	0.115
	Freq/gene (Targets)	**Freq/gene (Non-Targets)**					
QFM	1.170	1.074	—	—	$p < 0.0001$	$p < 0.0001$	—

Table 4.1: Statistics for motif analysis

targets (targets = 8.036, non-targets = 8.483). High enrichment was seen for the GAC motif (targets = 16.787, non-targets = 16.066), GACR (targets = 7.905, non-targets = 7.444), and GACARG motifs (targets = 0.875, non-targets = 0.773). While the WGGA frequency occurred a high frequency/kB (freq./kB = 16.531) there was significant difference in motif frequency compared to non-targets (freq./kB = 16.666). Finally, the QFM motif showed a significant enriched motif frequency per gene compared to non-targets (targets = 1.170, non-targets = 1.074). Importantly, the effect size (Cohen's *d*) is small in all cases indicating that although there is a significant difference in motif frequency/kB between targets (n = 2696) and non-targets (n = 11222) for all motifs except WGGA, the relationship may have moderately low practical significance. However, similar effect values were reported by Suhl et al. (Suhl et al., 2014). Finally, ρ was relatively small for all motifs (mean $\rho_{targets}$ = 0.371, mean $\rho_{non\text{-}targets}$ = 0.370) but in all cases where there was significant enrichment, the motif occurred with a greater frequency then we would expect by chance. The GAC motif showed the strongest motif frequency compared to chance ($\rho_{targets}$ = 0.955, $\rho_{non\text{-}targets}$ = 0.936). Overall, the results of our analysis were similar to the those seen in the analysis of the FMRP consensus data sets by Suhl et al., (Suhl et al., 2014) though there was some variability in the exact motif frequency in our data compared to the consensus data sets. Taken together, these results support the validity of our methodology and confirm that the motif frequency in the mRNA we isolated have similar enrichment over non-targets to previously identified FMRP targets.

Analysis of FMRP mRNA targets reveals differential binding and GO enrichment of trans-synaptic signaling genes between treatments

We have already seen that the application of Ro or Ro+Rapamycin leads to differential binding of FMRP mRNA targets (Figure 4.4E). Next, we sought to further

characterize the distributions of fold-changes for the FMRP targets in the two fold-change distributions under consideration (either Saline/Ro or Ro+Rapa/Ro). We predicted that if binding or release to FMRP is mTOR dependent then the administration of rapamycin will cause mRNA associated with specific biological processes to change simultaneously. To see if the fold-changes of any mRNA changed together within each of the different distributions we used univariate K-means clustering and determined the most representative GO biological processes found in each major cluster (k = 8). In both distributions (Saline/Ro or Ro+Rapa/Ro) we found that the trans-synaptic signaling genes were the most enriched GO biological process (Figure 4.5B, red bar). 60% of the trans-synaptic signaling genes whose fold changes clustered together in the Saline/Ro distribution also clustered together in the Ro+Rapa/Ro distribution. Furthermore, most of these mRNA had a positive fold-change value in both distributions which we would expect if binding or release was mTOR dependent.

We next visualized the intersection of the distributions described above (either $\log_2$(Saline/Ro) or $\log_2$(Ro+Rapa/Ro)) in order to find mRNA targets in both groups that changed in same direction (Figure 4.5C). Remarkably, we found that different FMRP targets were both released and bound in an mTOR-dependent manner from the mRNA population of FMRP targets that are translated when Ro activates mTOR ($\log_2$(Ro+Rapa/Ro) > 0) or those targets that are repressed when mTOR is on ($\log_2$(Ro+Rapa/Ro) < 0). We next used bivariate K-means clustering on the data this data, dividing it into $k = 11$ clusters (Figure 4.5C). Interestingly, the trans-synaptic genes group into different clusters suggesting that mTOR activity changes alters the expression of these genes by more than one mechanism.

The population of mRNA found in saline relative to those in Ro indicate either FMRP targets translated with Ro induction ($\log_2$(Saline/Ro) > 0) or those targets not

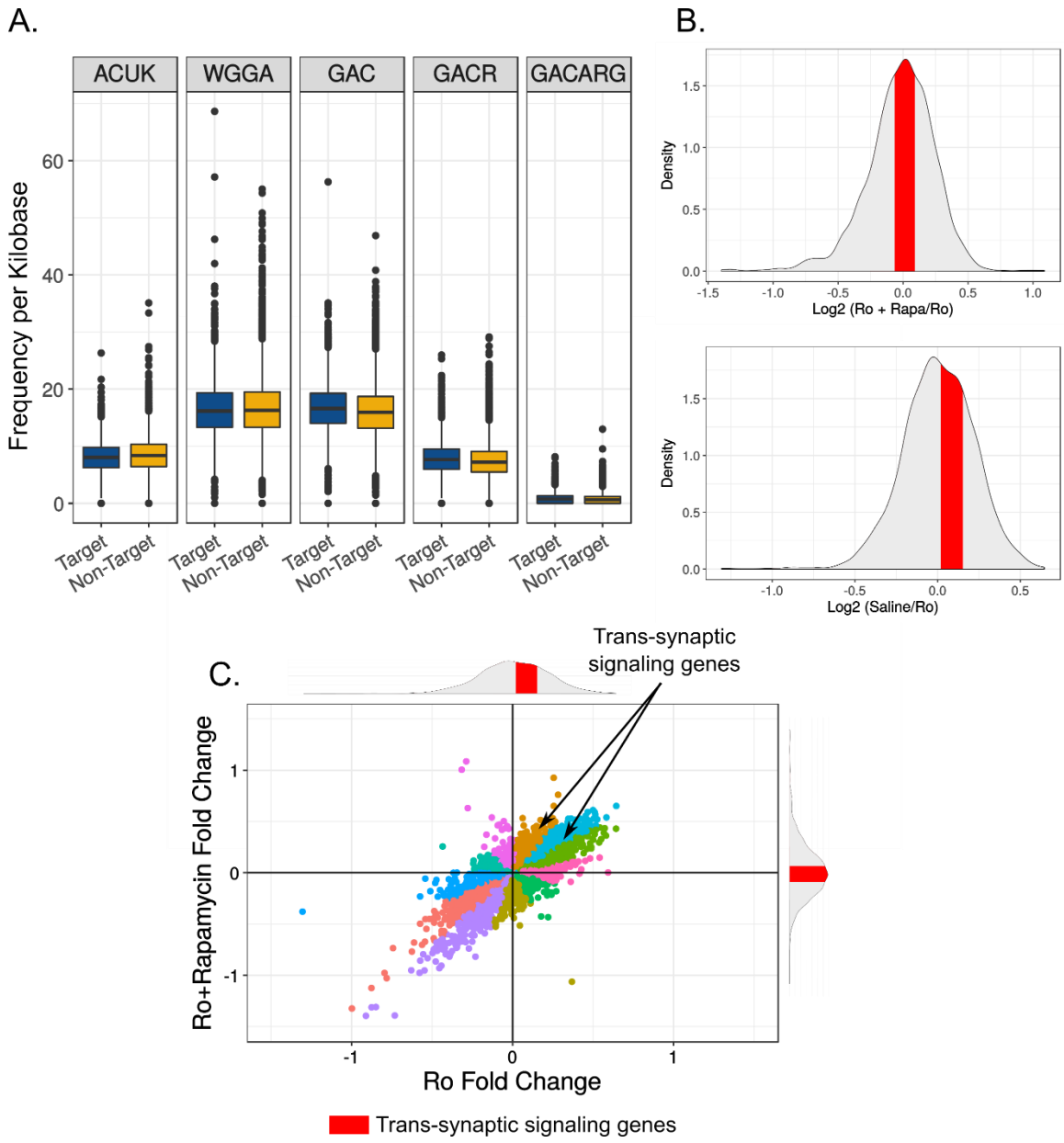


Figure 4.5: Motif analysis and clustering analysis for RIP-seq libraries

A. Boxplot representing data summarized in Table 4.1. **B.** Distribution of $\log_2$ transformed counts for either the Ro+Rapamycin/Ro fold-changes or Saline/Ro fold-changes. Red bar indicates cluster identified by univariate K-means that is significantly enriched with trans-synaptic signaling genes by DAVID GO biological process clustering. **C.** Scatterplot showing the $\log_2$(Ro+Rapaymcin/Ro) fold-changes on one axis and $\log_2$(Saline/Ro fold-change) on the other axis. Colored dots indicated various clusters identified by bivariate K-means clustering. Distribution (*top* and *right*) are the same as those shown in **C.**

translated with Ro induction (log_2(Saline/Ro) < 0) while the population of mRNA found in Ro+Rapa relative to Ro indicate FMRP targets translated when Ro turns mTOR on or FMRP targets repressed when mTOR is on. To further characterize the distributions of mRNA differentially binding to FMRP between drug treatments, we performed GO analysis to determine the most representative biological process clusters enriched for these mRNA populations (either Saline/Ro or Ro+Rapa/Ro) (Figure 4.6A,B). For either distribution, we first used the DAVID functional annotation chart to determine enriched GO biological processes for genes with a log_2 fold-change > 0 or a log_2 fold-change < 0 separately. We then computed the symmetric difference between these GO terms to find those that were uniquely expressed in either direction. Since this list was quite large, we utilized the same approach that we applied to the GO terms from the MDD genes mentioned earlier (Figure 4.8A) to condense the most enriched terms into categories via unsupervised hierarchical clustering. Surprisingly, the most enriched biological processes for targets translated with Ro induction (log2(Saline/Ro) > 0) or translated when Ro turns mTOR on (log_2(Ro+Rapa/Ro) < 0) were those under the parent GO category of GO:0099537: trans-synaptic signaling ($p < 0.0001$ for both fold-changes) (Figure 4.6A,B). These trans-synaptic genes mostly fall within the top right quadrant of the intersection of both of these fold-change distributions (targets translated when mTOR is on) indicated the mRNA involved in trans-synaptic signaling may be responsible for the synaptic remodeling seen in rapid antidepressant treatment (Figure 4.6C). Extending this analysis further we also clustered the trans-synaptic signaling genes found in either the Saline/Ro fold-change category (Figure 4.6D) or the Ro+Rapa/Ro fold-change category (Figure 4.6E). Interestingly, we see that the trans-synaptic signaling genes in the Saline/Ro fold-change distribution are more likely to change bidirectionally than do those found in the Ro+Rapa/Ro condition.

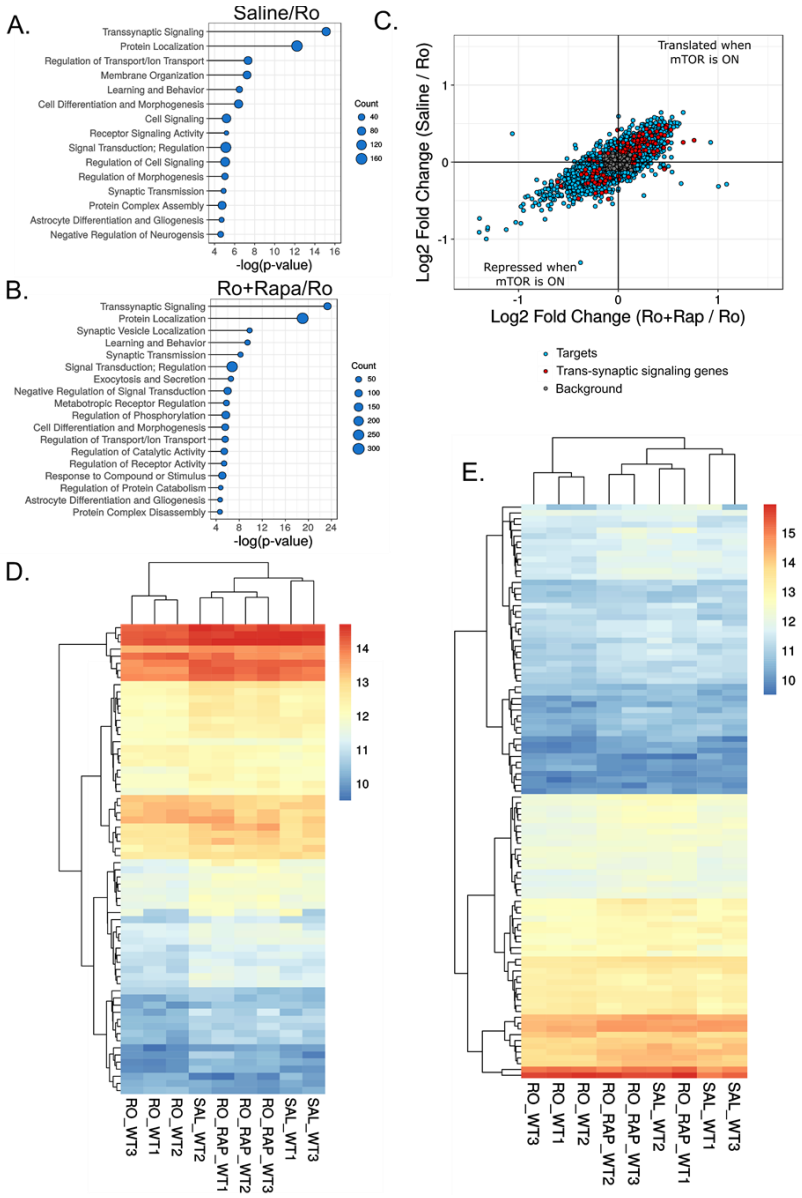


Figure 4.6: Identification and clustering of trans-synaptic signaling genes

A. Uniquely enriched GO biological process clusters for where $\log_2$(Saline/Ro) > 0. **B**. Uniquely enriched GO biological process clusters for where $\log_2$(Ro+Rapamycin/Ro) > 0. GO analysis for **A.** and **B.** performed with DAVID functional annotation chart. Chart processing is described in detailed in "Experimental Procedures". **C.** Scatterplot showing either $\log_2$(Saline/Ro) or $\log_2$(Ro+Rapamycin/Ro). Trans-synaptic signaling genes shown in red. **D.** Heatmap for trans-synaptic signaling genes found in $\log_2$(Saline/Ro). **E.** Heatmap for trans-synaptic signaling genes found in $\log_2$(Ro+Rapamycin /Ro). Raw count data was transformed using the regularized log function. Clustering using the average linkage method and Euclidean distance. Scales for **C.** and **D.** represent rlog-transformed raw counts

DISCUSSION

Trans-synaptic signaling proteins provide a potential target for the treatment of MDD

MDD is great economic burden in modern society and it is projected to be the 2nd cause of disability worldwide by 2020 (Lopez et al., 2006; Mathers et al., 2006). Numerous pharmacological interventions have been attempted but the most common approaches, SSRIs, have been shown to have low effectiveness and response rates (Turner et al., 2008; Fournier et al., 2010; Undurraga and Baldessarini, 2012; Greenberg et al., 2015). Furthermore, scales and criteria for diagnosing MDD suffer from a high heterogeneity of symptoms which is a major impediment in the design of successful drug treatments (Lohoff, 2010; Lux and Kendler, 2010; Goldberg, 2011; Hybels et al., 2011; de Vos et al., 2015; Lieblich et al., 2015; Fried, 2017). Recently, alternative approaches such as rapidly-acting antidepressants that act on NMDA receptors have proven to be a reliable approach to treating MDD due to their fast action and long-lasting effects (Dwyer and Duman, 2013). In this work, we have attempted to "reverse engineer" MDD by examining the transcriptome-wide changes that occur in the cortex in response to administration of the rapid antidepressant Ro-25-6981. The combined approach of high-throughput assays and bioinformatics analysis has identified numerous targets of the Fragile-X Mental Retardation protein that bind differentially in response to treatment with Ro. Among the targets of FMRP are mRNA whose protein products are involved in trans-synaptic signaling. Our work suggests that these proteins are responsible for the remodeling of the synapse seen upon administration of rapid antidepressants and provide a unique molecular view of both the etiology and treatment of MDD. Finally, our pharmacological approach

with CMap suggests a means of using drug discovery to find alternative drugs to treat MDD. Further investigation into the role of the trans-synaptic signaling proteins and their role in increasing activity of the presynaptic neuron ("presynaptic engagement") is currently underway. To date, only one work has examined the role of mTORC1 in this process (Henry et al., 2012) and our work underscores the importance of this pathway in retrograde signaling and presynaptic engagement processes.

A model for presynaptic engagement of mTOR-dependent FMRP targets

Under normal conditions, NMDA receptor activity is required for the activation of mTOR (Figure 4.7A, *left*). Our previous work has shown that NMDA antagonism results in the production of new $GABA_BR$ surface expression resulting from the de-repression of FMRP and the decoupling of $GABA_BR$ and GIRK channels due to increased stability of 14-3-3η (Workman et al., 2013; Workman et al., 2015; Wolfe et al., 2016). The FMRP targets that we recovered in our RIP-seq experiment from this phase are induced by Ro when mTOR is OFF. (Figure 4.7A, *center*). Our previous work also demonstrates that new $GABA_BR$ associates with L-Type Ca^{2+} channels to activate mTOR (Workman et al., 2013; Workman et al., 2015). The mRNA in this phase are induced by Ro when mTOR is ON (Figure 4.7A, *right*). The identification of trans-synaptic signaling proteins in our RIP-seq experiment suggests that the released FMRP targets undergo retrograde transmission to the presynaptic neuron to engage in processes of presynaptic formation to strengthen connectivity with the postsynaptic neuron (Figure 4.7A, *right*).

Our RIP-seq experiment has attempted to capture mRNA targets of FMRP from multiple phases of this process. We had predicted that when $\log_2(\text{Saline/Ro}) > 0$ or $\log_2(\text{Ro+Rapa/Ro}) > 0$, we are capturing a Ro-induced population that is either mTOR

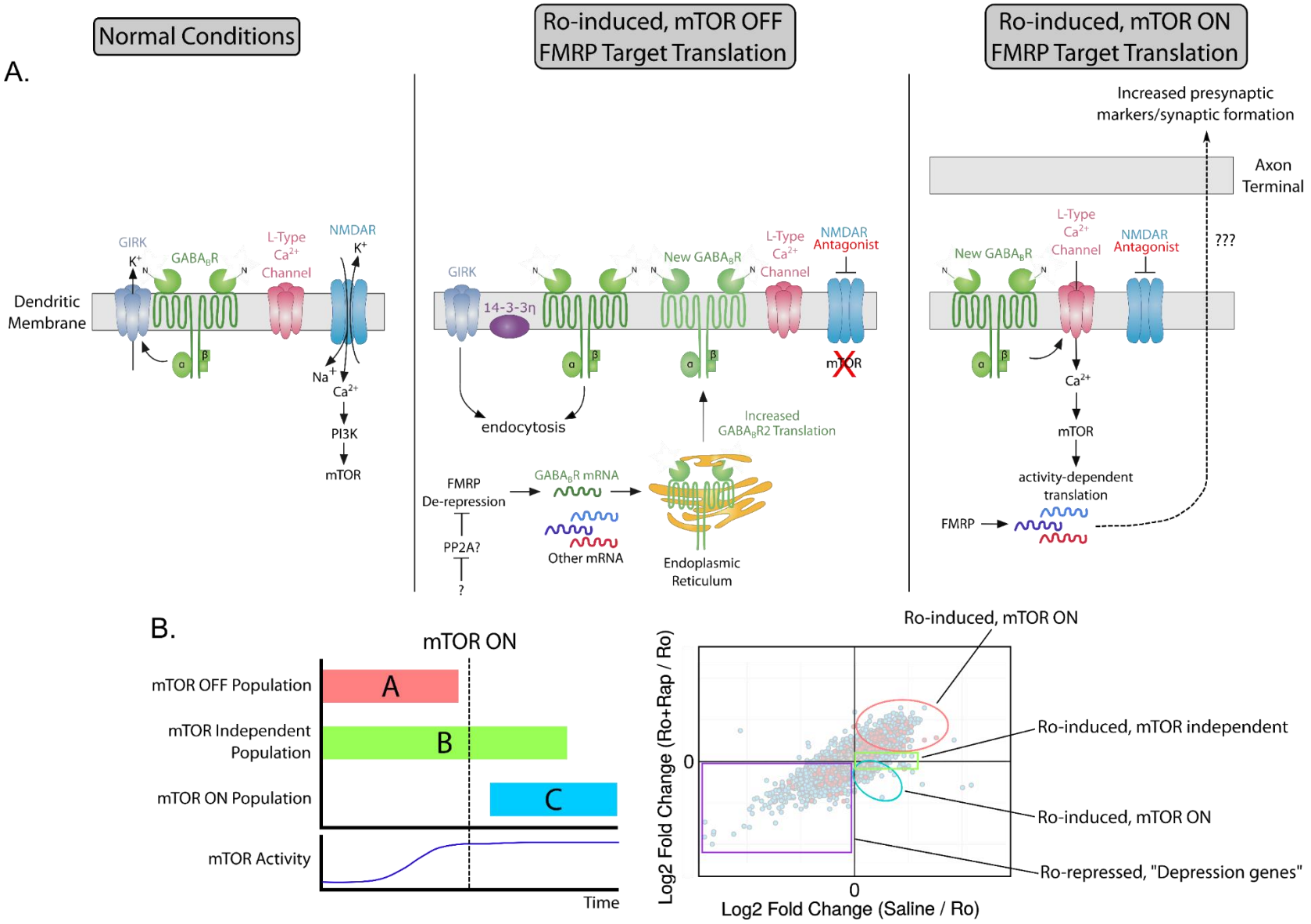


Figure 4.7: Model for mTOR-mediated synaptic remodeling through FMRP targets

independent or mTOR ON respectively (Figure 4.7B, *left*). Furthermore, the population where $\log_2$(Saline/Ro) < 0 or $\log_2$(Ro+Rapa/Ro) < 0 is predicted to capture a Ro-repressed population. Out data suggests that there are three separate populations of FMRP targets that are separately activated based on the activity state of mTOR (Figure 4.7B). mRNA present in the "A" phase represent the mTOR off population which corresponds to released FMRP targets as shown in Figure 4.7A, *center*. Among these targets are a basal population "B" whose fold-change is altered minimally by mTOR. We call these targets mTOR independent. Targets in "A" are predicted to include mRNA whose protein products activate mTOR. As mTOR activity increases (denoted by the hypothetical line graph for mTOR activity in Figure 4.7B, *left*), the mRNA in population "A" are rebound by FMRP and their fold-changes drop. Finally, the data suggests that there is the "C" population which includes targets activated by mTOR. These three populations correspond to mRNA

A. *Left*: Under normal conditions, glutamate binding to NMDA receptors on dendritic spines leads to the influx of Ca^{2+} which leads to the downstream activation of mTOR. *Center*: Under NMDA receptor antagonism, such as through a rapidly-acting antidepressant like Ro-25-6981, the first wave of FMRP target translation is initiated. Our data shows that GIRK channels are decoupled from $GABA_B$ receptors due to increased stability of 14-3-3η. Furthermore, this process also results in the de-repression of FMRP (possibly due to inhibition of PP2A) so that its target mRNAs are no longer bound and are thus available for translation. Among these are $GABA_B$ mRNA. Our past data indicates an increase in $GABA_B$R2 protein levels and surface levels of GABA receptors (Workman et al., 2013; Workman et al., 2015; Wolfe et al., 2016). *Right*: Another wave of translation is initiated after the newly synthesized $GABA_B$ receptors associate with L-Type Ca2+ channels. The influx of Ca2+ activates mTOR and translated targets no longer bound to FMRP. In this work, we identify some of these targets which include mRNA involved in trans-synaptic signaling. The final part of our model hypothesized that the protein products of these mRNA undergo retrograde signaling to the presynaptic cell to initiate remodeling processes that strengthen the presynapse ("presynaptic engagement") and postsynapse. **B.** RIP-seq experiment predicts three separate populations of FMRP targets involved Ro-induced mediation of depression. *Left*: Hypothetical depiction of distributions of each FMRP target population in time. Hypothetical depiction of mTOR activity is shown below the main plot. *Right*: Correlation scatterplot, derived from Figure 4.6A, shows the spread of our targets in the top-right and bottom-left quadrants. Colored boxes or ellipses correspond to populations identified in the *left* portion of Figure 4.6B. Additionally, a Ro-repressed population is show in the purple (Figure 4.6B, lower left quadrant) representing predicted FMRP target mRNA involved in MDD.

whose fold-changes can be visualized as a scatterplot of the two distributions $\log_2$(Saline/Ro) or $\log_2$(Ro+Rapa/Ro) where the top-right and bottom-right quadrants contain mRNA that are Ro-induced (Figure 4.7B, *right*). Visualizing the data in this way also suggests that there is another population of mRNA that is Ro-repressed (Figure 4.7B, *right, bottom-left quadrant*). We hypothesize that these mRNA are involved in MDD and must be actively repressed in for antidepressant efficacy. Finally, our data suggests that the trans-synaptic mRNA found in the top-right quadrant (Ro-induced and mTOR ON) is responsible for the remolding of the synapse. Collectively, our work provides a foundation for understanding the mechanistic processes behind the synaptic remodeling seen upon administration of rapidly-acting antidepressants. Our future work aims to investigate these processes in greater detail.

SUPPLEMENTAL FIGURES

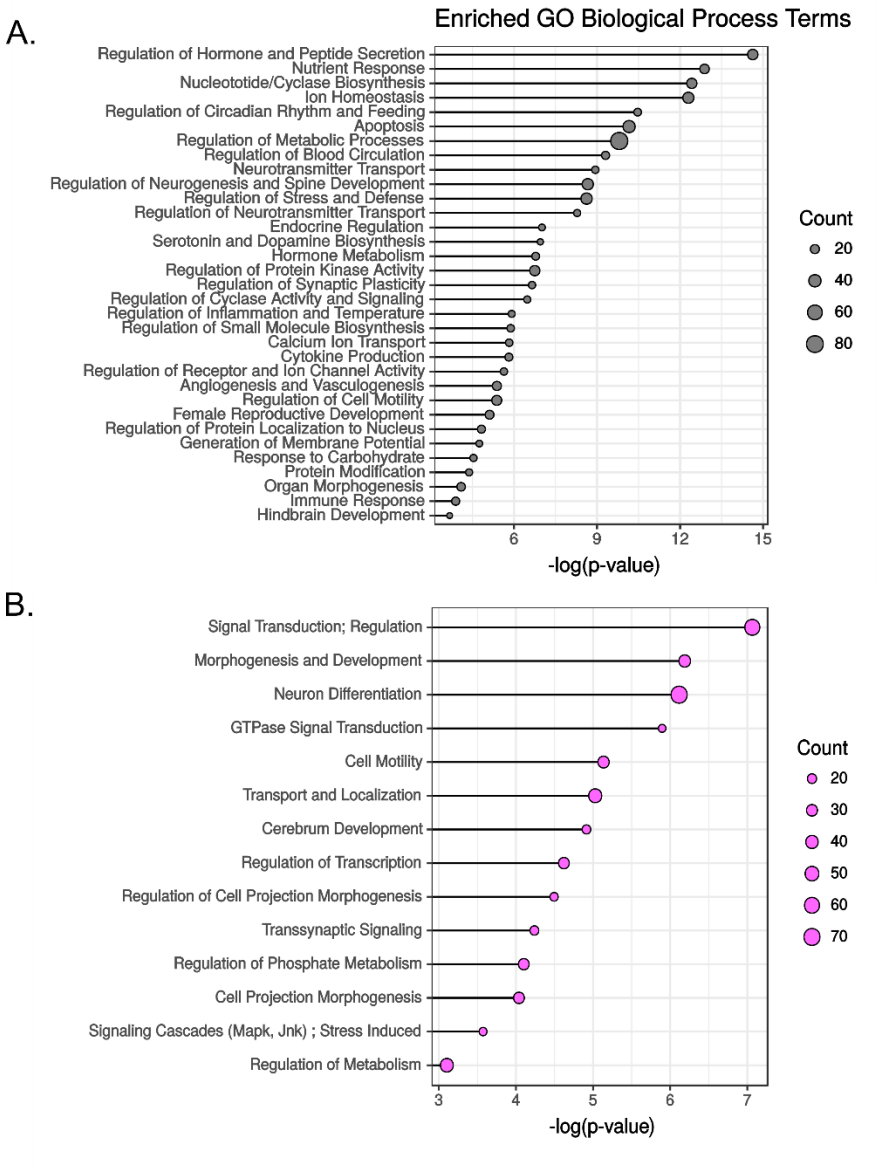


Figure 4.8: GO enrichment of biological process terms for MDD genes and FMRP consensus genes

Top GO terms for **A.** an MDD depression gene list derived from a literature search or **B.** genes from the Brown et al. and Darnell et al. data sets (Brown et al., 2001; Darnell et al., 2011) from the DAVID functional annotation chart. Terms from the chart were clustered by the degree of overlap of genes in each term to produce a condensed list of major biological processes. DAVID functional annotation charts were filtered for a p-value cutoff of 0.001 and an FDR cutoff of 0.05. p-values or counts shown in the chart are means for all terms within a specific named category along the y-axis. See "Experimental Procedures" for details.

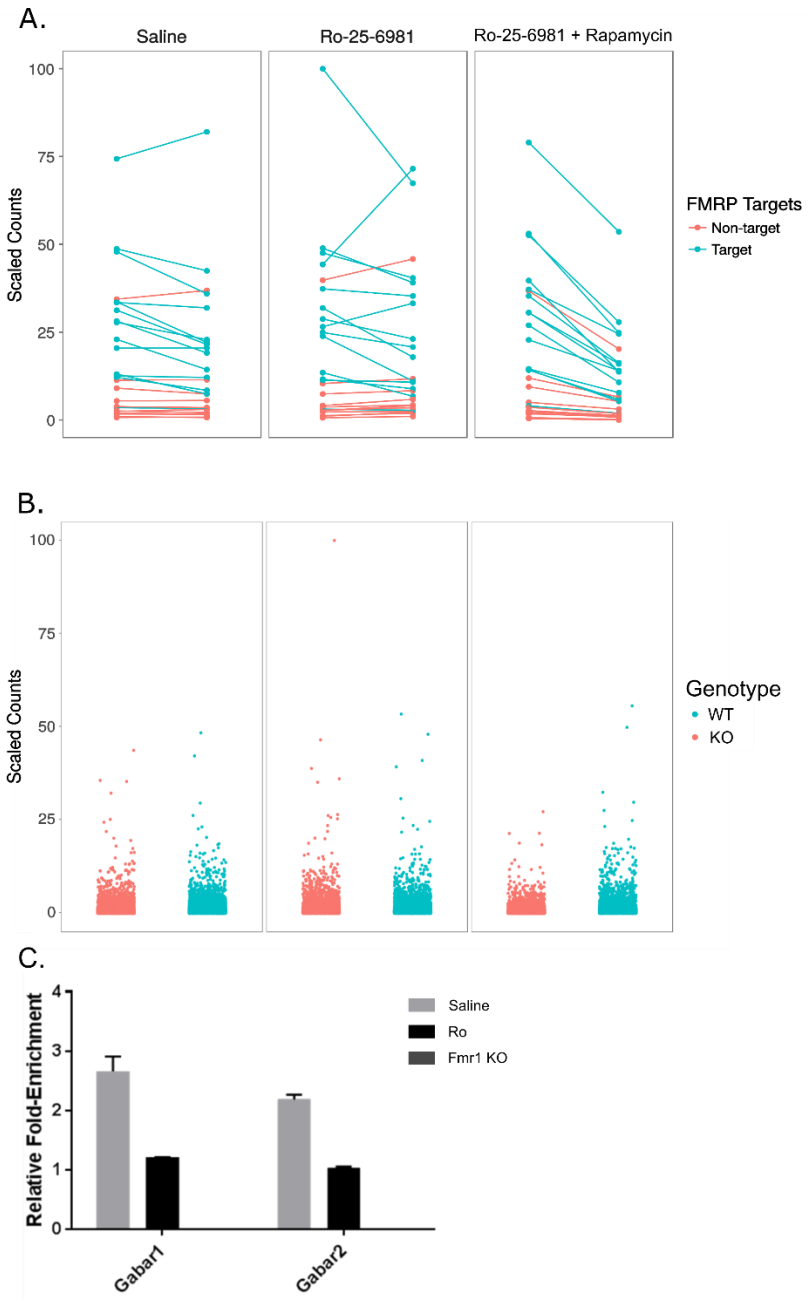


Figure 4.9: RIP-seq quality control

All counts shown in plots were scaled from 0-100 to facilitate comparisons. **A.** Normalized counts for FMRP targets and non-targets (list derived from (Darnell et al., 2011)) across treatments. **B.** Normalized counts of entire RIP-seq libraries (either WT or KO). **C.** Relative fold-enrichment as determined by real-time qPCR relative to 10% input control ($\Delta\Delta Ct=2^{-(Ct\ FMRP\ RIP-Ct\ IgG\ RIP)-(Ct\ FMRP\ input-Ct\ IgG\ input)}$). FMRP binds to $GABA_BR1$ and $GABA_BR2$ in Saline or Ro-25-6981 treated animals but not in *Fmr1* KO animals from RNA derived from RIP of FMRP.

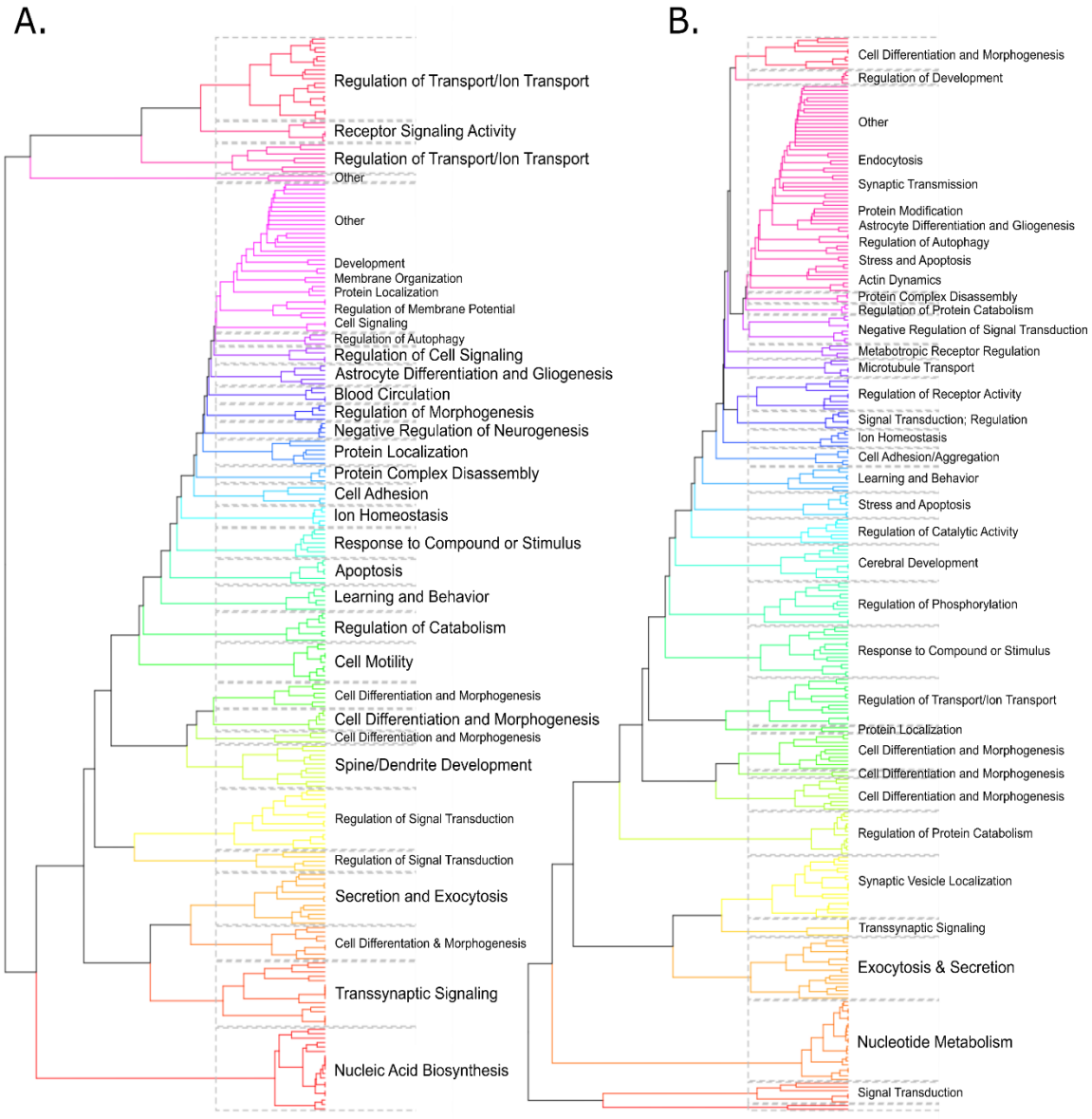


Figure 4.10: Hierarchical clustering of overlapped GO terms derived from DAVID GO biological process function annotation chart

GO terms unique to either **A.** Saline/Ro > 1 or **B.** Ro+Rapamycin/Ro > 1 were derived from the DAVID GO biological process functional annotation chart. A p-value cutoff of 0.001 and an FDR cutoff of 0.05 was used to filter the table. The resulting terms were then clustered by creating a similarity matrix based on the number of overlapping genes in each GO term. Clustering was performed using the average linkage method and used Euclidean distance. These terms were then used to create the charts seen in Figure 4.6 A,B. See "Experimental Procedures" for details.

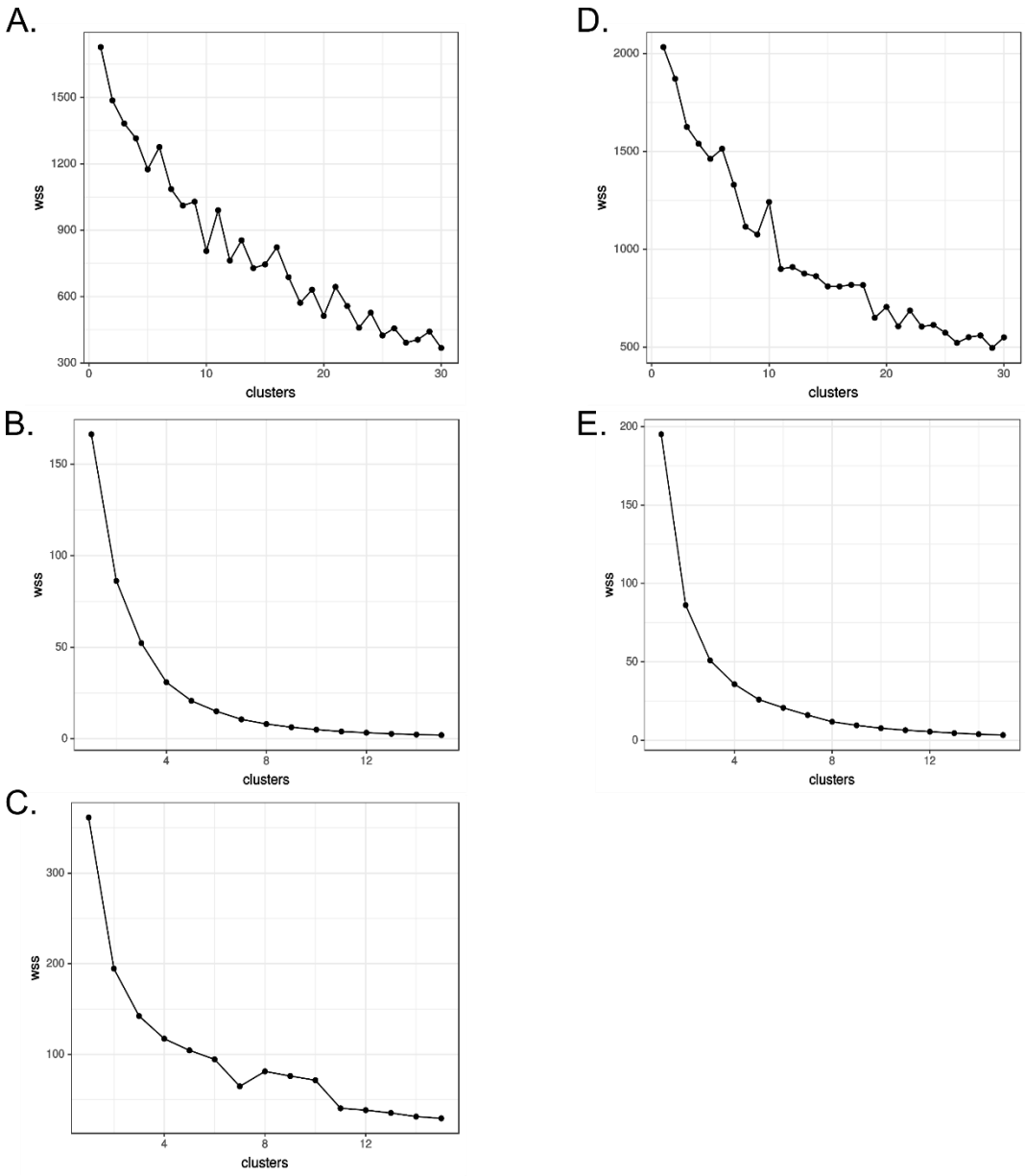


Figure 4.11: Within-sum-of-squares (WSS) estimation for number of cluster used in clustering algorithms

WSS was used to determine a suggested number of clusters to divide our data wherever a clustering algorithm (univariate K-means, bivariate K-means, or hierarchical) was to be applied. **A.** WSS calculation of clusters for Saline/Ro clustering (Figure 4.6A). **B.** WSS calculation of clusters for Ro+Rapamycin/Ro clustering (Figure 4.6B). **C.** WSS calculation of clusters for univariate K-means of Saline/Ro distribution (Figure 4.5B, *top*). **D.** WSS calculation of clusters for univariate K-means of Ro+Rapamycin/Ro clustering (Figure 4.5B, *bottom*). **E.** WSS calculation of clusters for bivariate K-means of both Saline/Ro and Ro+Rapamycin/Ro distributions (Figure 4.5C).

SUPPLEMENTAL TABLES

RNA mass recovered (ng)						
Treatment	**Wild-type (Replicate)**			**Knockout (replicate)**		
	1	2	3	1	2	3
Ro-25-6981	36	67	48	3	5	3
Ro-25-6981 + Rapamycin	56	101	75	47	48	102
Saline	54	326	139	29	37	35
RNA integrity number (RIN)						
Ro-25-6981	9.0	8.6	9.1	1.0	8.4	1.1
Ro-25-6981 + Rapamycin	8.6	8.8	8.8	9.0	7.2	8.9
Saline	8.6	8.3	6.7	8.7	8.8	8.3

Table 4.2: RNA mass recovered and RIN values for RNA immunoprecipitation of FMRP targets

Library	Length (nt)	Input_Reads	Mapping_Rate	Multiple_Alignments
RO_KO2_T1	20-51	436,922	16.10%	15.60%
RO_KO2_T2	20-51	399,555	16.10%	15.50%
RO_KO2_T3	20-51	4,787,614	16.60%	15.00%
RO_KO2_T4	20-51	3,990,679	16.40%	15.10%
RO_RAP_KO1_T1	20-51	8,253,643	72.90%	4.40%
RO_RAP_KO1_T2	20-51	8,140,014	73.00%	4.40%
RO_RAP_KO1_T3	20-51	9,524,091	73.50%	4.40%
RO_RAP_KO1_T4	20-51	8,983,390	73.20%	4.40%
RO_RAP_KO2_T1	20-51	1,539,793	76.00%	3.60%
RO_RAP_KO2_T2	20-51	1,447,787	76.00%	3.60%
RO_RAP_KO2_T3	20-51	5,235,945	75.40%	3.60%
RO_RAP_KO2_T4	20-51	4,919,303	75.30%	3.60%
RO_RAP_KO3_T1	20-51	8,735,516	73.30%	4.00%
RO_RAP_KO3_T2	20-51	8,640,851	73.30%	4.00%
RO_RAP_KO3_T3	20-51	230,114	73.60%	4.00%
RO_RAP_KO3_T4	20-51	9,691,348	73.30%	4.00%
RO_RAP_WT1_T1	20-51	9,425,580	81.90%	3.90%
RO_RAP_WT1_T2	20-51	9,329,940	82.00%	3.90%
RO_RAP_WT1_T3	20-51	3,634,272	82.50%	3.90%
RO_RAP_WT1_T4	20-51	2,678,561	82.30%	3.90%
RO_RAP_WT2_T1	20-51	9,389,927	86.50%	4.30%
RO_RAP_WT2_T2	20-51	9,330,571	86.60%	4.30%
RO_RAP_WT2_T3	20-51	3,847,304	86.70%	4.30%
RO_RAP_WT2_T4	20-51	2,873,233	86.60%	4.30%
RO_RAP_WT3_T1	20-51	7,040,228	86.40%	4.10%
RO_RAP_WT3_T2	20-51	7,030,839	86.50%	4.10%
RO_RAP_WT3_T3	20-51	1,221,288	86.70%	4.10%
RO_RAP_WT3_T4	20-51	504,767	86.70%	4.10%
RO_WT1_T1	20-51	8,568,455	39.00%	5.10%
RO_WT1_T2	20-51	8,499,710	39.00%	5.00%
RO_WT1_T3	20-51	5,282,085	40.10%	5.00%
RO_WT1_T4	20-51	4,971,366	39.90%	5.00%
RO_WT2_T1	20-51	973,892	60.30%	5.60%
RO_WT2_T2	20-51	915,550	60.40%	5.60%
RO_WT2_T3	20-51	9,052,870	61.30%	5.50%
RO_WT2_T4	20-51	8,498,721	61.00%	5.60%
RO_WT3_T1	20-51	7,787,301	55.40%	5.00%
RO_WT3_T2	20-51	7,826,769	55.40%	5.00%
RO_WT3_T3	20-51	5,785,648	56.40%	4.90%
RO_WT3_T4	20-51	4,770,198	56.10%	4.90%
SAL_KO1_T2	20-51	3,360,188	62.30%	4.30%
SAL_KO1_T1	20-51	3,511,776	62.20%	4.40%
SAL_KO1_T3	20-51	975,474	63.20%	4.30%
SAL_KO1_T4	20-51	329,254	63.00%	4.30%
SAL_KO2_T1	20-51	7,887,843	46.10%	4.70%
SAL_KO2_T2	20-51	7,762,342	46.00%	4.70%
SAL_KO2_T3	20-51	6,240,194	47.10%	4.60%
SAL_KO2_T4	20-51	5,855,520	46.90%	4.60%
SAL_KO3_T1	20-51	141,624	55.60%	4.40%
SAL_KO3_T2	20-51	38,776	55.70%	4.40%
SAL_KO3_T3	20-51	8,361,756	56.10%	4.30%
SAL_KO3_T4	20-51	7,898,366	55.70%	4.30%
SAL_WT1_T1	20-51	7,366,745	46.40%	4.60%
SAL_WT1_T2	20-51	7,395,787	46.50%	4.60%
SAL_WT1_T3	20-51	4,801,547	47.60%	4.60%
SAL_WT1_T4	20-51	3,949,727	47.20%	4.60%
SAL_WT2_T1	20-51	7,165,746	74.50%	3.90%
SAL_WT2_T2	20-51	7,629,193	74.40%	3.90%
SAL_WT2_T3	20-51	961,778	74.70%	3.90%
SAL_WT2_T4	20-51	208,678	74.50%	3.90%
SAL_WT3_T1	20-51	6,961,446	87.40%	4.20%
SAL_WT3_T2	20-51	6,952,007	87.50%	4.20%
SAL_WT3_T3	20-51	3,979,856	87.70%	4.20%
SAL_WT3_T4	20-51	2,940,141	87.60%	4.20%

Table 4.3: RIP-seq library and alignment data

Chapter 5: Conclusion

In the last decade, we have seen a vast expansion in the application of high-throughput experiments in providing models for the physiological relevance of locally synthesized proteins. This expansion includes the earliest microarray studies on synaptic fractions, where it was considered fantasy to suggest ~400 mRNAs are localized in dendrites, to large scale RNA sequencing data suggesting closer to 2500 mRNAs reside in dendrites. Collectively, these data are expanding our knowledge of how cellular processes of learning and memory occur. Protein synthesis-dependent synaptic plasticity has been described by many investigators utilizing cellular physiology and well characterized protein synthesis inhibitors; however mechanistic details, at the molecular level, have lagged behind. Currently, many of the missing details are being provided by high-throughput experiments that analyze the changes in the proteome or transcriptome in response to synaptic stimuli. However, the sudden interest in these types of experiments is a direct result of (1) the decreased cost of high-throughput experiments, (2) the availability and decreased cost of high-performance computing tools for processing large amounts of high-dimensional data, (3) the increasing accuracy and sensitivity of high-throughput sequencing technologies, and (4) the widespread development of biological assays that allow researchers to rapidly capture the interactions and dynamics of thousands of molecules simultaneously. Many other fields have embraced this technological shift and have begun to apply high-throughput technologies to many different avenues of research. The field of local dendritic translation could greatly benefit from these techniques as neuronal functioning and connectivity is due to a very complicated and dynamic framework of tightly-controlled interactions between the genome, proteome, and transcriptome.

The studies outlined here (Chapters 2-4) show the potential of high-throughput methodologies combined with bioinformatics to transform our knowledge of the molecular underpinnings of processes of learning and memory as well as neuronal dysfunction. Our approach in Chapter 2 used mass spectrometry to reveal that mTORC1 controls the expression of numerous proteins bidirectionally. We then explored these results in diseases of overactive mTOR to demonstrate that many proteins associated with these diseases change bidirectionally. This provides new avenues of exploration to determine the function of these proteins in situations where overactive mTOR represses their expression rather than actives it. In Chapter 3, we continued to explore the relationship between diseases of overactive mTOR through a number of bioinformatics approaches. We identified PARK7 as a unique protein hub in an interaction network perhaps linking them together under the common theme of dysregulated translation. We also identified a novel function for PARK7 in TSC and provided data suggesting possible pharmacological approaches to treating disorders involving mTOR and PARK7 targets. Finally, in Chapter 4, we used a high-throughput RIP-seq approach to understand the dynamics of FMRP target binding in the presence of a rapidly-acting antidepressant. This experiment provides the first data that attempts to understand the mechanism of this class of antidepressants in MDD, suggesting a more general mechanism by which MDD occurs in the brain. Overall, our combined bioinformatics and high-throughput experiment approaches illustrates the vast potential to answer many long-standing questions in the field of local dendritic translation.

Firstly, though the stages of LTP and synaptic plasticity are broadly understood, the specific dynamics underlying each step are not known and many different models with slight variation are published each year (Manninen et al., 2010). Elucidation of protein targets and protein-RNA interactions at different stages will provide further insight into the regulation and dynamics of these stages. Furthermore, LTP is typically studied at a single

synapse when in reality memory formation likely underlies many different processes, including homeostatic plasticity and synaptogenesis, at multiple synapses and systems of neurons taken as a whole. Some attempts have been made at providing a unified model of LTP (Michmizos et al., 2011) and many of the missing details could be provided by high-throughput experiments that analyze the change in the proteome or transcriptome in response to synaptic stimuli. Secondly, the structural and morphological changes that occur in response to stimulated synapses is still poorly understood. Some important questions include: How do locally translated transcripts contribute to spine growth and synaptic remodeling? How are these processes controlled? How do signals from stimulated synapses lead to large-scale restructuring of synaptic connectivity between neurons and neuronal networks? Thirdly, and important but often overlooked feature involved in long-term memory formation is presynaptic engagement and retrograde signaling. Finally, many neurological disorders may be the consequence of a mutation in a single protein but this can have far-reaching effects on many other networks and processes. High-throughput experiments provide a powerful tool to show the alteration of molecular networks that occur as a result of disease states.

There have been ongoing efforts in the last decade to provide rule-based meta-languages to accurately model biological networks including Kappa (Feret et al., 2009) and NFSim (Sneddon et al., 2011). Such computational modeling systems effectively represent complex biological networks as a rule-based system and provide powerful prediction and simulation capabilities. A rule-based approach has already been used to characterize the structural dynamics of the PSD (Sorokina et al., 2011). This approach provides a more accurate and representative view of interactions between proteins since it accounts for stoichiometric properties within the protein-protein interaction network that reflect the dynamic nature inherent to molecular systems. Another possibility for representation and

prediction in biological networks is the use of Bayesian model-inference approaches (Pe'er, 2005; Klinke, 2009). As data continues to increase from high-throughput experiments in the coming years, this information will become further integrated into computational modeling of the signaling networks providing a more unified view of the processes that underlie synaptic plasticity, memory, and the coordination of these effects through local translation. As techniques are refined and technology improves in the coming decade we foresee the frequent application of high-throughput technology to provide both novel candidates and coordinated local expression of proteins that make individual synapses plastic and unique.

Appendix A: Pushing the threshold: How NMDAR antagonists induce homeostasis through protein synthesis to remedy depression[10]

ENGAGING HOMEOSTATIC MECHANISMS WITH PHARMACOLOGICAL INTERVENTION TO TREAT MAJOR DEPRESSIVE DISORDER

The idea of using modern medicine to push "out-of-balance" or diseased neurons to their tipping point seems outrageous. However, basic scientists do so on a regular basis. Experimental paradigms that induce long-lasting changes in synaptic efficacy—in one direction—are used to study cellular mechanisms of learning and memory (synaptic plasticity). Sustained extremes in neuronal activity, inherent to synaptic plasticity, eventually cause neurons and neuronal circuits to become unstable (Davis and Goodman, 1998; Turrigiano and Nelson, 2000; Pozo and Goda, 2010). To avoid this problem, healthy neurons adjust their synaptic strength to maintain optimal action potential firing rates. This process is known as homeostatic plasticity and requires dynamic molecular changes (Turrigiano, 2008).

Depression, anxiety disorders, and addiction, result from imbalanced neuronal networks (Figure A.1) (McClung and Nestler, 2008). Such perturbations in neuronal signaling, while severe, may not tip the scale enough for homeostatic mechanisms to kick in. Can modern medicine capitalize on what basic science has taught us about homeostasis? In other words, can pharmacological intervention in vivo push neuronal activity to its limit to induce homeostatic response that will eventually cause neurons and neuronal circuits to self-correct? Studies on rapid antidepressant therapies (i.e. N-methyl-D-aspartate receptor (NMDAR) antagonists argue yes (Figure A.1B-C).

[10] Originally published in (Raab-Graham, K.F., Workman, E.R., Namjoshi, S., Niere, F. (2016). Pushing the threshold: How NMDAR antagonists induce homeostasis through protein synthesis to remedy depression. Brain Res. 1647: 94-104. doi: 10.1016/j.brainres.2016.04.020. I wrote the "Regulation of Autophagy" subsection and created Table A.1. figures A.2, and figure A.3.

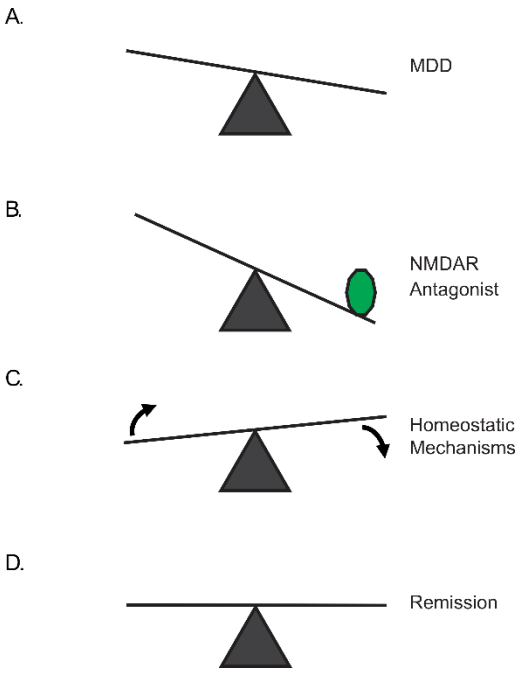


Figure A.1: Model for rapid antidepressant activation of homeostatic mechanisms

A. MDD leads to a network that is outside of the ideal homeostatic range, as indicated by the out of balance scale. **B.** NMDAR antagonists (green circle) pushes the network further outside the ideal range activating homeostatic mechanisms **C.** that result in relief from MDD and remission **D.**

Are new antidepression therapies necessary?

Major Depressive Disorder (MDD) is a severe neuropathophysiology that will affect up to 17% of the population at some point during their lifetimes (Zarate et al., 2010). The most common pharmacological therapies, selective serotonin reuptake inhibitors (SSRIs), target the serotonergic system by blocking serotonin uptake. Unfortunately, only ~37% of individuals afflicted with MDD experience relief from their depressive symptoms with SSRIs (Murrough, 2012). Moreover, with continued use of SSRIs, only ~ 35-50% will go into remission (Rush et al., 2006; Trivedi et al., 2006; Trivedi et al., 2008; Murrough and Charney, 2012). Thus, finding new effective treatments for MDD is imperative. New therapies, such as rapid antidepressants and deep brain stimulation, are hypothesized to work by engaging network and cellular homeostatic mechanisms. Consequently, these therapies that are revolutionizing depression treatment, are predicted to drive the neurons to self-correct.

What are Rapid Antidepressants?

Rapid antidepressants are antagonists of NMDARs (Table A.1). NMDAR is a ligand and voltage-gated cation channel that is best characterized for its role in shaping synaptic strength during synaptic plasticity (Blanke and VanDongen, 2009). Behavioral assessment has shown that ketamine, Ro 25-6981 and other FDA approved NMDAR antagonists have remarkable efficacy in reversing treatment-resistant depression phenotypes. Additionally, NMDAR antagonists increase synaptogenesis, suggesting that increased neuronal communication arises from NMDAR blockade (Li et al., 2010; Ibrahim

et al., 2011; Lepack et al., 2014). While Trullas and Skolnick observed 25 years ago that NMDAR antagonists have antidepressant properties in rodents, recent clinical studies on the efficacy of ketamine and other NMDAR antagonists have renewed the field's interest

Mechanism	NMDAR antagonist	Brain region	Effect	References
Plasticity	AP5	HPC	Induces $GABA_BR$ shift from opening K channels to facilitating L-type calcium channel activity	(Workman et al., 2013)
		HPC	Potentiates AMPAR-mediated evoked neurotransmission at rest	(Nosyreva et al., 2013)
	Ketamine	HPC	Low-dose ketamine promotes low-dose AMPA-induced synaptogenesis (synapsin)	(Akinfiresoye and Tizabi, 2013)
		PFC	Increases of spine number	(Li et al., 2010)
		PFC	Increases serotonin-induced EPSC	(Li et al., 2010)
	MK-801	PFC	Switches HPC-induced LTD to LTP	(Thomas et al., 2014)
	Ro-25-6981	HPC	Rescues LTP and gamma oscillation in LTP-deficient Tn65DN mouse	(Hanson et al., 2013)
		HPC	Impairs gamma oscillation in normal mouse	(Hanson et al., 2013)
Structural plasticity	Ketamine	PFC	Promotes spine growth	(Li et al., 2011)
Intrinsic excitability	AP5	HPC cultures	14-3-3 eta decouples $GABA_BRs$ from GIRK channels	(Workman et al., 2015)
	Ketamine	Cortex	Inhibits in I_h current in layer V cortical pyramidal cells	(Chen et al., 2009)
		HPC	Knockdown of HCN1 in the dorsal hippocampal produced antidepressant-like behaviors in the FST, similar to ketamine injected rats	(Kim et al., 2012)
Transcription	AP5	HPC cultures	Increases expression of HDAC4 mRNA and protein	(de Bartolomeis et al., 2013)
	Ketamine	IC	Increases expression Homer1a mRNA	(de Bartolomeis et al., 2013)
		AC, SS	Increases expression of Arc mRNA	(de Bartolomeis et al., 2013)
		CP	Decreases expression of PSD95 mRNA	(de Bartolomeis et al., 2013)
		MC	Decreases expression of Homer1b mRNA	(de Bartolomeis et al., 2013)
	MK-801	CP	Decreases expression of PSD95 mRNA	(de Bartolomeis et al., 2013)
		HPC astrocyte cultures	Increases expression of BDNF, GFAP and TrkB mRNAs	(Yu et al., 2015)

Table A.1: Homeostatic response to NMDAR antagonism

Mechanism	NMDAR antagonist	Brain region	Effect	References
Translation (eEF2-dependent/global	Ketamine	HPC, FC	Increases expression of BDNF	(Autry et al., 2011)
	MK-801	HPC	Increases expression of BDNF	(Autry et al., 2011)
		HPC, FC	Increases expression of Arc	(Autry et al., 2011)
Translation (mTOR-dependent)	AP5	HPC cultures	Baclofen promotes expression of BDNF	(Workman et al., 2013)
	Ketamine	PFC	Increases expression of synaptic proteins (GluA1, PSD95, Synapsin 1)	(Li et al., 2010)
Translation (unknown mechanism)	AP5	HPC cultures	Increases expression of $GABA_BR2$ protein	(Workman et al., 2015)
	Ro-25-6981	HPC	Increases expression of $GABA_BR2$ protein	(Workman et al., 2015)
Apoptosis	AP5	HPC cultures	Increases expression of pro-apoptotic genes expression	(Chen et al., 2009)
Autophagy	Ro-25-6981	CTX	Delays the expression of TBI-induced autophagic proteins	(Bigford et al., 2009)
		Neuronal cultures	Inhibits glutamate excitotoxicity-induced cell death	(Bigford et al., 2009)

Table A.1: Homeostatic response to NMDAR antagonism (continued)

in these drugs as potential antidepressants (Zarate et al., 2006; Skolnick et al., 2009). Still, at the molecular level, it is uncertain as to why blocking NMDARs relieves depressive symptoms. This review focuses on the convergence of the clinical application of NMDAR antagonists and the basic science of neuronal homeostasis focusing on protein synthesis-dependent pathways.

NMDAR ANTAGONISM LEADS TO HOMEOSTATIC PLASTICITY

A neuron employs two forms of homeostatic responses—global and local synaptic scaling—to deal with extreme high or low activity (Turrigiano, 2012). Global synaptic scaling induces a slow, cell-wide compensation in excitatory and inhibitory receptors throughout the soma and dendritic arbor. In contrast, local synaptic scaling occurs within hours as a response to local activity fluctuations, allowing dendrites or groups of synapses to restore operation at an optimal range (Queenan et al., 2012).

NMDAR subunit composition underlies the homeostatic mechanism of Metaplasticity

Metaplasticity is a form of homeostasis describing how cells or circuits reset their baseline activity levels in response to their previous activity (Abraham et al., 2001). Metaplasticity has been studied mostly at the circuit level, albeit it has been reported to occur at a single synapse (Lee et al., 2010a). For example, in seminal studies that deprive rodents of light during a critical period of visual developmental, the absence of visual input alters the threshold for synaptic plasticity in the forms of long-term potentiation (LTP) and long-term depression (LTD) (Kirkwood et al., 1996; Philpot et al., 2003). At the molecular level, light deprivation induces cell-wide forms of synaptic scaling by upregulating

excitatory receptors and downregulating inhibitory receptors. Such a shift in the ratio of excitatory and inhibitory inputs will favor excitation and is consistent with a reduced threshold to induce LTP (Turrigiano, 2012). Interestingly, upon a short exposure to light, causing cortical activity to rise, more excitatory synaptic input is required to induce LTP (Philpot et al., 2003). These experiments, as well as many that have come before and after, serve as the basis for metaplasticity (Bear, 1995).

At the molecular level, changes in NMDAR subunit composition that regulate calcium dynamics likewise support metaplasticity (Yashiro and Philpot, 2008). NMDARs are obligate heteromultimers consisting of an NR1 subunit in combination with NR2A-D and NR3A-B (Glasgow et al., 2015). NR2A and B are the most critical subunits in setting the membrane threshold (Yashiro and Philpot, 2008). NR2B-containing NMDARs are slower to deactivate and thus carry more calcium per unit current relative to NR2A-containing receptors (Sobczyk et al., 2005; Glasgow et al., 2015). A decrease in the NR2A/2B ratio, an adjustment that appears to be seemingly modest in nature, significantly slides the threshold to favor LTP. As expected, NR2B-containing channels are upregulated relative to NR2A in the dark; and upon activation, allow for greater influx of calcium in spines (Yashiro and Philpot, 2008). Thus, one would predict that blocking NR2B activity under conditions of reduced cortical activity would provide the biggest push toward silencing synaptic activity. Furthermore, such a push may be sufficient to engage alternative homeostatic mechanisms to return the neuron back to an optimal operating range.

MAMMALIAN TARGET OF RAPAMYCIN (MTOR) SERVES AS A THERMOSTAT IN NMDAR ANTAGONIST-INDUCED HOMEOSTATIC RESPONSE

mTOR senses the functional, operating range of a neuron

How does one identify if a neuron is "out-of-range"? Is there a molecular marker whose activity level gauges the general health of a neuron? In MDD, it has been suggested that mTOR activity levels may serve as a molecular thermostat as it can gauge the efficacy of antidepressant drugs (Zarate et al., 2013). mTOR is a serine/threonine kinase that controls protein synthesis, cell growth and metabolism by sensing the levels of nutrients, growth factors, cellular energy, and stress. mTOR is conserved from yeast to human, signifying its critical importance in cellular homeostasis. mTOR consists of two complexes, mTORC1 and C2. mTORC1 regulates protein synthesis through phosphorylation of two proteins that regulate mRNA translation—eukaryotic initiation factor 4 binding protein (eIF4BP) and ribosomal S6 kinase (Hay and Sonenberg, 2004). In the brain, mTOR mediates dendritic protein synthesis, critical for synaptic plasticity and memory consolidation (Graber et al., 2013b).

In many diseases, mTORC1 is often overactive or underactive, consistent with a neuron operating outside its optimal range. For example, in epilepsy, Alzheimer's disease, and autism spectrum disorders—including Fragile X Syndrome and Tuberous Sclerosis Complex—mTORC1 signaling is excessive (Pei and Hugon, 2008; Zeng et al., 2009; Gross et al., 2010; Sharma et al., 2010; Spilman et al., 2010; Santini and Klann, 2011; Zeng et al., 2011; Sahin, 2012; Crino, 2015; Sosanya et al., 2015a). Furthermore, in MDD and Parkinson's disease, mTORC1 signaling is reduced (Jernigan et al., 2011; Fragkouli and Doxakis, 2014). It is currently unknown how mTORC1 gets stuck and fails to return to basal levels in these diseases. What can be done to "unstick" mTOR activity, allowing it to return to basal levels? Insights from NMDAR antagonists used for treating MDD may offer

an answer. Several labs have shown that mTORC1 activity increases with a single treatment of NMDAR antagonists in vivo. The increase in mTORC1 is required for the antidepressant efficacy of NMDAR antagonists (Li et al., 2010; Workman et al., 2013). In the brain, where the circuitry is intact and the release of neurotransmitters is difficult to control, deciphering how NMDAR blockade increases mTOR activity is challenging. Utilizing a simplified model of cultured hippocampal neurons, where one has tight control over the activation of neurotransmitter receptors, blockade of NMDARs with (2R)-amino-5-phosphonovaleric acid (AP5) reduces basal mTOR activity (Workman et al., 2013). These results indeed suggest that reducing NMDAR signaling pushes mTORC1 activity levels down and would be predicted to induce homeostasis (Figure A.1). However, as mTORC1 activity remains low, the trigger for the cell to push back and return mTORC1 activity to control levels is absent in these cultures. Collectively, these data provide the first hint that an extrinsic factor, such as a neurotransmitter, is necessary to restore mTORC1 activity to normal basal levels.

NMDAR activity at rest shapes the ability of metabotropic GABABR signaling to turn mTORC1 activity up.

What turns mTOR activity on with NMDAR blockade? One clue came from studies showing that NMDARs and γ-aminobutyric acid B receptors (GABABR), that are metabotropic G-protein-coupled receptor, extensively interact in a feedback loop (Guetg et al., 2010; Terunuma et al., 2010; Workman et al., 2013). GABA is considered as the major inhibitory neurotransmitter in the brain (Jembrek and Vlainic, 2015). Under normal physiological conditions, GABABRs inhibit neuronal activity through pre- and postsynaptic mechanisms. Postsynaptically, GABABRs mediate the slow inhibitory postsynaptic potentials via the activation of postsynaptic, G-protein-regulated, inwardly

rectifying potassium (GIRK) channels (Padgett and Slesinger, 2010). Stimulating NMDARs in cultured hippocampal neurons leads to endocytosis of GABABRs (Guetg et al., 2010; Terunuma et al., 2010). Likewise, blocking NMDAR signaling in vitro increases the dendritic surface expression of GABABRs (Workman et al., 2013). This reciprocal relationship between synaptic activity and GABABR surface expression, at first glance, would appear to provide positive feedback—less inhibition with synaptic stimulation and more inhibition with reduced synaptic activity. Surprisingly, NMDAR blockade causes GABABR function to shift from signaling that opens GIRK channels to signaling that increases dendritic Ca2+ requiring L-type Ca2+ channel activity (Workman et al., 2013). This increase in Ca2+, in turn, activates mTORC1 in primary apical dendrites that are enriched in GABAergic synapses (Megias et al., 2001; Workman et al., 2013). Moreover, GABABR signaling through mTORC1 promotes the synthesis of the brain derived neurotropic factor (BDNF), which is required for new synapse formation with rapid antidepressants (Figure A.2) (Li et al., 2010; Workman et al., 2013). These results emphasize that restoring neuronal communication leads to the sustained antidepressive efficacy of NMDAR antagonists (Li et al., 2010; Li et al., 2011; Liu et al., 2012).

RAPID ANTIDEPRESSANT EFFICACY REQUIRES DISTINCT STAGES OF NEURONAL PROTEIN SYNTHESIS

Evoked versus miniature NMDAR currents

One perplexing concept regarding the effectiveness of NMDAR antagonists as rapid antidepressant is the requirement that the membrane be depolarized for NMDARs to open. NMDARs require the binding of two neurotransmitters, glutamate and glycine. Even in the presence of these molecules the channel remains closed due to Mg^{2+} blocking the

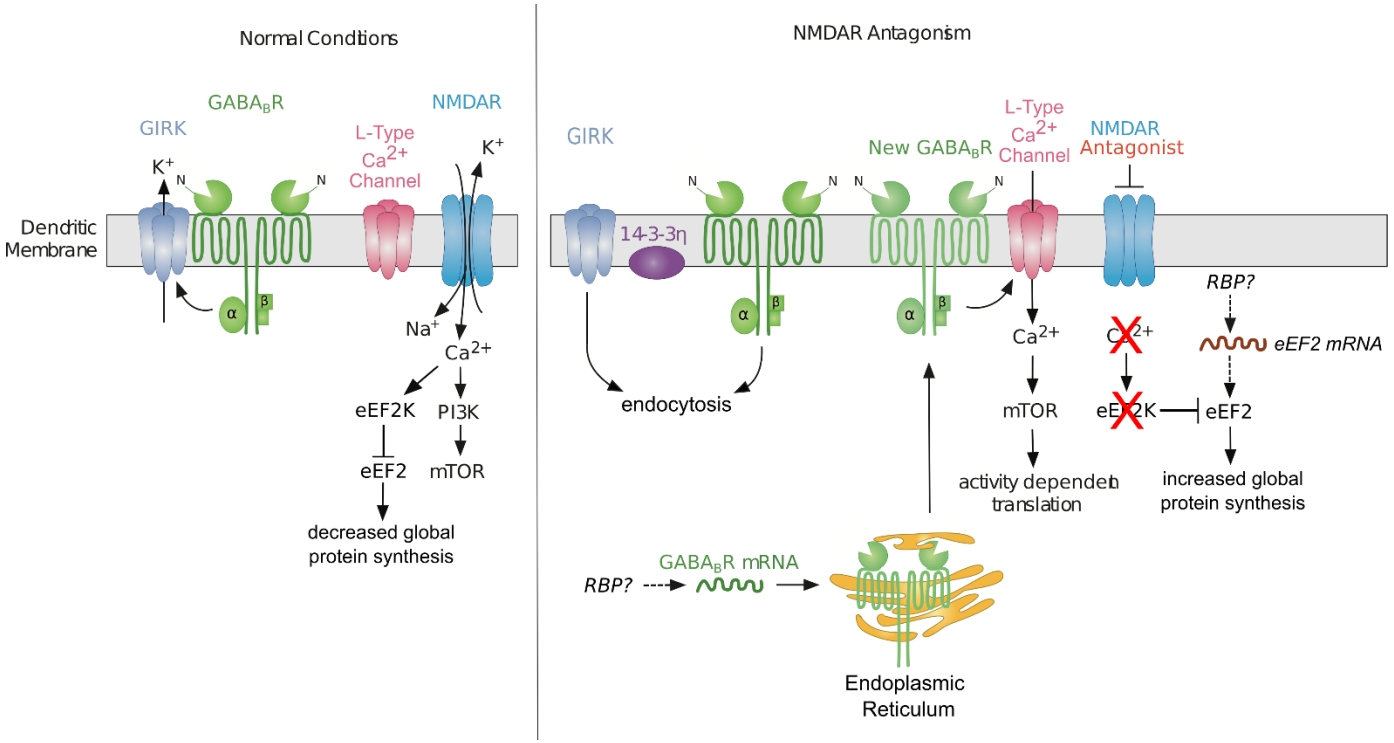


Figure A.2: Cartoon describing molecular pathways triggered by NMDAR antagonists

Left panel: Activation of GABA$_B$Rs signal to open GIRK channels reducing dendritic calcium levels. Calcium entry through NMDARs activates the translation elongation factor kinase eEF2K that in turn phosphorylates eEF2, inhibiting global translation. *Right panel*: In the presence of NMDAR antagonists GABA$_B$R signaling shift to facilitating L-type calcium channel activity and activates mTOR-dependent protein synthesis. This requires the protein synthesis and insertion of new GABA$_B$Rs into the dendritic membrane. What RNA binding protein (RBP) mediates GABA$_B$R mRNA repression (*left*) and or translation (*right*) is an open question. In addition, reduced NMDAR signaling to eEF2K increases the activity of eEF2 allowing for global translation. eEF2 protein expression increases when mTOR activity is reduced. If new eEF2 mRNA translation and the RBPs that regulate its expression is unknown.

ion-conducting pathway. Depolarization of the membrane dislodges Mg^{2+}, and facilitates the movement of cations, particularly Ca^{2+}, in and out of the cell (Burnashev et al., 1992). Considering the strict biophysical properties of the channel, one would predict that NMDAR antagonists quiet active neurons and neuronal circuits. However, some studies suggest that the Mg^{2+} block of NMDARs is partial at the resting membrane potential (Reese and Kavalali, 2015). Fortunately, since NMDAR antagonists come in many flavors, pharmacological approaches can dissect NMDAR resting current blockade from evoked and total NMDAR blockade. For example, MK-801, which selectively blocks open NMDARs, and memantine, only blocks NMDARs when Mg^{2+} is dispelled from the pore conducting pathway, are widely used. It is generally considered that MK-801 blocks both evoked and resting NMDAR currents, whereas memantine only blocks NMDARs engaged in firing Remarkably, rodents treated with MK-801 display improvement on behavioral despairs, while with memantine they do not (Autry et al., 2011; Gideons et al., 2014). Thus, NMDAR antagonists that are used as rapid antidepressants (e.g. MK-801) are effective at reducing the activity of neurons engaged in firing, as well as a population of NMDARs that are open at rest via spontaneous activity.

NMDAR blockade-induced translation initiates two phases of antidepressant efficacy

Studies suggest that antidepressant efficacy requires two phases—induction and sustained. While it is clear that mTORC1-dependent protein synthesis is required for the long-lasting effects of rapid antidepressants, the molecular changes that initiate the process remain unclear. Does the initiation phase require protein synthesis? Work on local homeostatic synaptic scaling demonstrates that acute blockade of spontaneous NMDAR

activity increases the miniature excitatory synaptic current (mEPSC) mediated by the glutamate receptor that contains α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) 1 subtype (GluA1). Such AMPAR-mediated synaptic scaling requires de novo dendritic protein synthesis of GluA1 (Sutton et al., 2006).

Notably, several groups have observed increased expression of GluA1 with administration of NMDAR antagonists *in vivo* (Li et al., 2010; Autry et al., 2011; Nosyreva et al., 2013; Workman et al., 2013). A separate set of studies have also demonstrated that blocking spontaneous NMDAR activity increases evoked AMPAR-mediated synaptic potentiation in the CA1 region of hippocampal slices. This form of potentiation was sustained and further amplified an hour after the antagonist had been washed away (Nosyreva et al., 2013). Interestingly in a BDNF knockout mouse, the initial synaptic potentiation due to NMDAR antagonists was present and similar to wildtype slices; however, the sustained potentiation was absent. In contrast, the inclusion of a protein synthesis inhibitor blocked the early and sustained potentiation induced by NMDAR blockade, indicating a requirement for protein synthesis in both plasticities (Nosyreva et al., 2013). These results confirm the necessity of BDNF for the long-lasting effects on synaptic potentiation due to NMDAR blockade and perhaps hint at a separate requirement for protein synthesis for the initial potentiation phase.

If mTORC1 activity mediates translation of mRNAs that are required for the sustained antidepressant effects of NMDAR antagonists, what mediates protein synthesis in the initiation phase? Schuman and colleagues proposed that Ca2+ entry through open NMDARs promotes the kinase activity of the eukaryotic elongation factor 2 kinase (eEF2K), which inhibits eEF2—an elongation factor that promotes global mRNA translation. As they had predicted, NMDAR blockade with AP5 in the presence of tetrodotoxin (TTX), which inhibits action potential, reduced Ca2+ entry, inactivated

eEF2K, and increased local protein synthesis. These results led the authors to conclude that NMDAR-mediated mEPSCs inhibit local protein synthesis (Sutton et al., 2004; Sutton et al., 2006; Sutton and Schuman, 2006). These studies, for the first time, assigned a function to spontaneous release and mEPSCs, as a critical regulator of homeostatic protein synthesis (Sutton et al., 2004). In light of MDD, do NMDAR-mediated mEPSCs occur more often or have a greater impact in "depressed" neurons or neuronal circuits? There is no direct experimental evidence that addresses this question. However, work on rapid antidepressants demonstrates that eEF2-dependent translation is upregulated with NMDAR antagonists in vivo. Furthermore, NMDAR-antagonist-induced antidepressant response can be bypassed completely by treating rodents with drugs that directly activate eEF2 (Autry et al., 2011).

NMDAR ANTAGONISTS INDUCE STRUCTURAL HOMEOSTASIS THAT CORRELATES WITH ANTIDEPRESSANT EFFICACY

One of the most striking effects of rapid antidepressants is increased spine density (Ohgi et al., 2015). Dendritic spines are the primary location of excitatory synapses in the adult brain. Spine shape, size, and number vary over the normal lifetime of an individual, and many neuropathologies display atypical spine distribution, shape, and number (Bourne and Harris, 2008). Structural homeostasis is defined by neuronal architecture changes in response to long-term fluctuations in activity (Yin and Yuan, 2014). Over the years, sensory paradigms that prompt homeostatic plasticity also induce morphological changes in spines (Butz et al., 2009; Keck et al., 2011). For example, sensory deprivation results in enlarged spine heads and increased postsynaptic AMPAR expression (Wallace and Bear, 2004; Keck et al., 2013). Moreover, whisker trimming paradigms that reduce sensory inputs to the cortex decreased spine elimination (Zuo et al., 2005).

In 1999, Kirov and Harris found that blocking NMDARs in an acute hippocampal slice preparation significantly increased the spine density in dendrites (Kirov and Harris, 1999). Duman and colleagues, a decade later, also reported that NMDAR antagonism increased dendritic spines *in vivo* by using mice that sparsely expressed yellow fluorescent protein in the prefrontal cortex (Li et al., 2010; Li et al., 2011). With new synapse formation one would expect increased pre- to postsynaptic signaling. As predicted, mEPSC frequency in layer II/III cortical neurons increased 24 hours post-rapid antidepressant injection (ketamine or Ro 25-6981) (Miller et al., 2014).

mTOR mediates NMDAR antagonist-induced spine formation

mTOR activity increases spine density. In neurons with hyperactive mTOR signaling, elevated spine number can be reduced with the mTORC1 inhibitor rapamycin (Tavazoie et al., 2005). It has been suggested that mTORC1-dependent protein synthesis is required for new spine formation and may be the key link for remission from depression (Li et al., 2010; Li et al., 2011). In support of this view, direct rapamycin infusion into the prefrontal cortex prevents the ketamine-induced increase in spine number (Li et al., 2010; Li et al., 2011). Of note, BDNF, which is translated via eEF2 and mTORC1, is necessary for the increased expression, maturation and stability of spines (Bennett and Lagopoulos, 2014). In light of how BDNF is translated, speculations abound that BDNF synthesis is required for the expression of the early and late phases of NMDAR antagonist-mediated antidepressant response, whereby the "early" synthesized BDNF is secreted to act upon postsynaptic TrkB receptors that bind BDNF with high affinity (Autry et al., 2011). Upon BDNF binding, TrkB signals cascade to turn mTOR on (Takei et al., 2004; Kavalali and Monteggia, 2012; 2015). BDNF secretion from the postsynaptic neuron is likely to require

L-type calcium activity, perhaps through the activation of GABABRs (Workman et al., 2013; Lepack et al., 2014). Further work is needed to test this hypothesis and to identify the intermediate events that lead to new synapse formation. In spite of incomplete mechanistic details, these data suggest that altering neuronal activity with natural stimuli such as light, experimentally with whisker trimming, or pharmacologically with NMDAR antagonists promotes changes in spine shape, size and density to maintain homeostasis.

EVIDENCE FOR ALTERNATIVE WAYS OF INDUCING HOMEOSTATIC PLASTICITY TO REMEDY DEPRESSION

Intrinsic Excitability

If NMDAR antagonists exert their action by lowering synaptic activity to engage homeostatic mechanisms, then directly altering ion channel function or density (intrinsic excitability) of the dendritic membrane may serve the same function as NMDAR antagonism. For example, knockdown of the HCN1 channel produces an antidepressant behavioral phenotype and increases mTOR and BDNF, two molecular mediators of rapid antidepressant efficacy (Kim et al., 2012). While ketamine is best known for its ability to block NMDARs, it has been shown to also block HCN channels (Chen et al., 2009). Moreover, as mentioned above, switches in receptor-channel signaling are involved in formation and reversal of depressive symptoms. Specifically, upon NMDAR antagonism, GABABR signaling to GIRK channels decreases while L-type channel activity increases. This switch in turn enhances mTOR activity and BDNF expression (Workman et al., 2013). Blocking the decoupling of GABABR from GIRK in the presence of NMDAR antagonists restores GABABR-mediated hyperpolarization and prevents rapid antidepressant efficacy in vivo (Workman et al., 2015). Interestingly, sleep deprivation also mediates a rapid

antidepressant effect (McClung, 2007). In fruit flies, a screen designed to find mutants with a short sleeping phenotype identified the shaker locus (Kv1) (Cirelli et al., 2005), a channel whose expression is repressed by mTORC1 activity (Raab-Graham et al., 2006; Sosanya et al., 2013; Sosanya et al., 2015a; Sosanya et al., 2015b; Niere et al., 2016). It will be interesting if like HCN, reduced levels of Kv1.1 feedback to increase mTORC1/BDNF signaling resulting in antidepressant behaviors. All in all, these findings suggest that changes in intrinsic excitability can regulate the mTOR/BDNF antidepressant pathway.

Regulation of Autophagy

Aberrant autophagy has been suggested to underlie depression (Abelaira et al., 2014). Autophagy is triggered when the cell is deprived of nutrients or placed in stressed conditions and serves to recycle organelles, clear protein aggregates, and eliminate non-essential macromolecules in order to aid in cell survival. In certain cases, cell signaling events that induce autophagy eventually lead to apoptosis (Marino et al., 2014). Tissue studies in human and animal models have shown that depression is characterized by the loss of neural and glial cells due to chronic stress (Banasr and Duman, 2008; Krishnan and Nestler, 2008; Pittenger and Duman, 2008). Various stress paradigms in rats have been shown to upregulate levels of pro-apoptotic proteins that are released through mitochondrial outer membrane permeablization as well as the release of caspase 3 (Bachis et al., 2008; Kosten et al., 2008). Furthermore, a transcriptome-wide profile of post-mortem human prefrontal cortex brain tissue from 14 pairs of subjects analyzed by microarray revealed the over-expression of apoptosis factors and inflammatory cytokines (Shelton et al., 2011). Notably, antidepressants have been shown to upregulate levels of anti-apoptotic proteins, particularly members of the Bcl-2 family involved in the suppression of

mitochondrial outer membrane permeabilization, a key step in the intrinsic apoptosis pathway (McKernan et al., 2009).

mTOR is known to control the homeostasis between cell growth (synthesis of new molecules) and autophagy (breakdown of current cell structures and macromolecules). mTOR is a negative regulator of autophagy but its mechanism of control is poorly understood. mTOR phosphorylation of ULK1 prevents activation of the ULK1 complex, a regulator of autophagy and inducer of apoptosis (Ganley et al., 2009; Jung et al., 2009; Lee and Tournier, 2011; Mukhopadhyay et al., 2015). Importantly, a reduction in both phosphorylation and expression of components of the mTOR signaling pathway has been observed in a postmortem examination of prefrontal cortex tissue obtained from depressed subjects in a series of studies. These proteins include NR2A, NR2B, mGluR5, PSD-95 as well as mTOR and its downstream targets involved in the control of translation p70SK6, eIF4E, p-eIF4E, eIF4B, and p-eIF4B (Feyissa et al., 2009; Deschwanden et al., 2011; Jernigan et al., 2011). All of this suggests an important role for mTOR in the homeostatic balance between autophagy and protein synthesis that may be relevant in the context of depression.

Finally, the NMDAR antagonists, 3-methyladenine (3-MA) and Ro-25-6981, have been shown to suppress the induction of autophagy and neuronal cell death (Bigford et al., 2009; Sadasivan et al., 2010; Xin et al., 2011). Bigford and colleagues demonstrated that Ro-25-6981 blocked the induction of autophagy and lowered the levels of autophagic proteins both in vitro and in vivo. This and previously discussed work suggest the possibility that the Ro-25-6981-mediated repression of NR2B signaling may by neuroprotective through the mTOR pathway. Thus, mTOR may be uniquely positioned (1) to promote the growth of new spines, (2) to regenerate dendritic architecture through protein synthesis, and (3) to suppress autophagy-mediated apoptosis in neurons injured by

chronic stress (Jia and Le, 2015). Future studies will be needed to determine the manner in which NR2B-specific antagonists are able to trigger these effects.

PERSPECTIVES AND OPEN QUESTIONS

Insight into depression and antidepressant therapies through the identification of proteins synthesized when mTORC1 is inhibited

Are there other ways to coax the compromised or diseased neuron to self-correct? Clues may come from recent work that examines dendritic protein synthesis in the presence of the mTORC1 inhibitor rapamycin in vivo (Niere et al., 2016). While the role of mTORC1 activity to promote translation of mRNAs is grounded in its established function of phosphorylating eIF4BP and ribosomal S6 kinase (Hay and Sonenberg, 2004), a few studies suggest a separate and likely just as important role for mTORC1 activity in repressing mRNA translation of specific transcripts (Raab-Graham et al., 2006; Sosanya et al., 2013; Niere et al., 2016). Interestingly, Niere et al., using the unbiased approach of mass spectrometry, demonstrates that the number of proteins that increase at the synapse with mTORC1 inhibition is approximately equal in number to those whose expression is reduced (Niere et al., 2016). Is it possible that some of these mRNAs are homeostatic in nature, and that the proteins they encode may predispose the neuron to turn mTORC1 back on? Notably, some of the proteins that are synthesized with brief mTORC1 inhibition overlap with proteins previously associated with depression and antidepressant properties (Figure A.3A-C) (Fatemi et al., 2001; Tadic et al., 2007; Dlugos et al., 2009; Unschuld et al., 2009; Lee et al., 2010b; Luo et al., 2010; Rietschel et al., 2010; Aragam et al., 2011; Galeotti and Ghelardini, 2011; Shyn et al., 2011; Dekker et al., 2012; Sakaida et al., 2013; Ray et al., 2014; Gatt et al., 2015; Gray et al., 2015; Ogawa and Kunugi, 2015; Skoog et

al., 2015; Wagner et al., 2015; Yoshimasu et al., 2015; De Vry et al., 2016; Liu et al., 2016; Stachowicz et al., 2016). For example, Homer 1a was found to be upregulated in several paradigms of antidepression therapy including sleep deprivation, electroconvulsive shock, ketamine, imipramine, and fluoxetine (Serchov et al., 2016). Moreover, the Homer1 knockout mouse displays phenotypes consistent with anxiety and depression (Szumlinski et al., 2005). To date, it is unclear as to what protein synthesis pathways and translation factors are utilized to translate these homeostatic mRNAs. One would predict that eEF2 is involved but may not be the only factor. Notably, eEF2 is a protein whose expression is differentially regulated by mTORC1 activity, in a subcellular specific manner (Figure A.3C). Is synthesis of eEF2 mRNA required to turn mTORC1 activity on with NMDAR blockade (Figure A.2)? Such a result would suggest that additional translational mechanisms upstream of eEF2 are required for antidepressant efficacy. Nevertheless, coordinated stages of protein synthesis utilizing different translational mechanisms such as eEF2 and mTORC1 are required for rapid antidepressant efficacy, and mTORC1-dependent translation of BDNF is critical for the sustained effects of NMDAR antagonists (Lepack et al., 2014).

The unexplored role of RNA-binding proteins in rapid antidepression therapy

Targeting new protein synthesis pathways as a means to treat depression is a relatively new idea. As noted above, protein synthesis of specific proteins is orchestrated to mediate short- and long-term effects. Spatial accumulation of mRNAs in dendrites is widely believed to be a mechanism that regulates temporal gene expression. Cis-acting elements such as noncoding RNAs and RNA binding proteins (RBP) reversibly bind to the

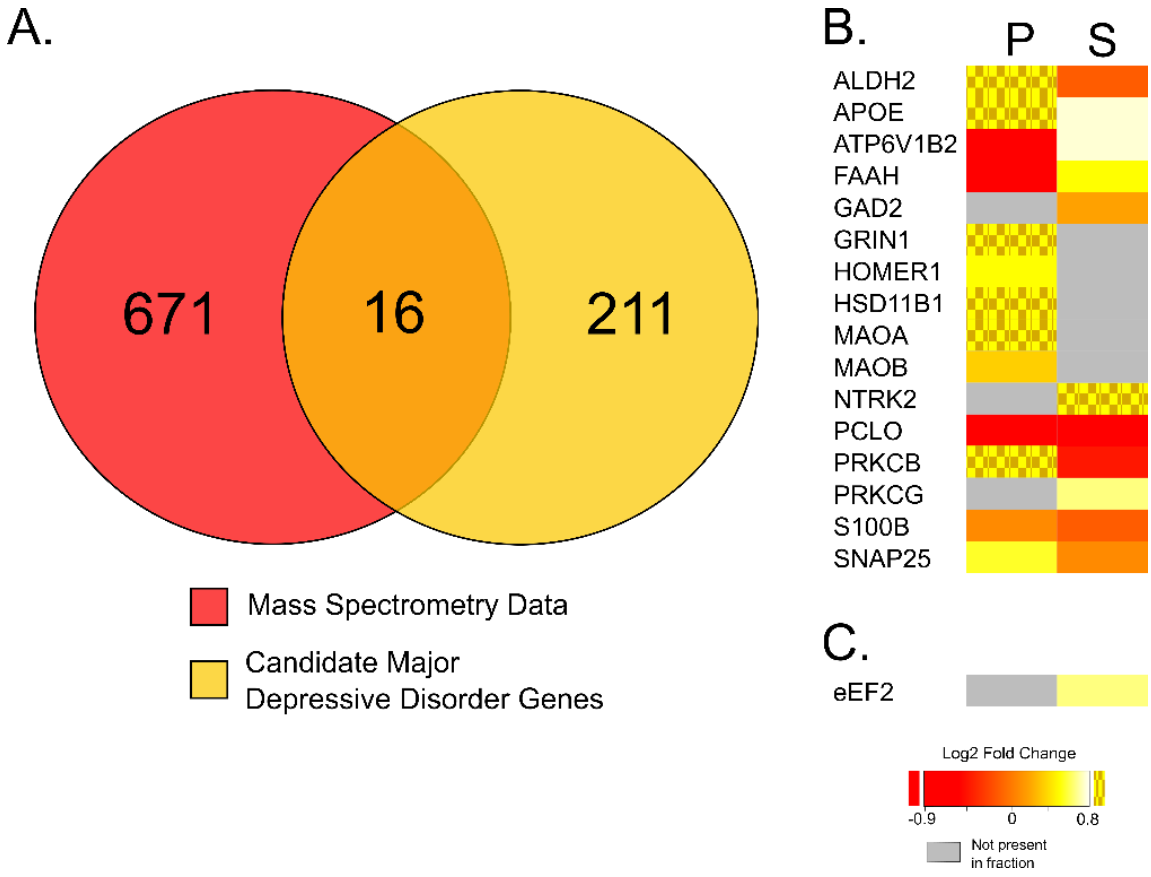


Figure A.3: Protein expression regulated by mTORC1 activity overlap with depression

A. Venn diagram illustrating several proteins identified by mass spectrometry of synapses with brief *in vivo* mTORC1 inhibition overlap with proteins associated with depression. **B.** Heat map of 16 overlapping proteins in the postsynaptic density fraction (P) and soluble fraction (S) of synaptoneurosomes isolated from cortices of control and rapamycin injected rats. Note, many of those proteins increase with mTORC1 inhibition (yellow). **C.** Heat map of EEF2, demonstrating increased expression in soluble fraction of synaptoneurosomes with brief mTORC1 inhibition. Data obtained from (Niere et al., 2016) and (Fatemi et al., 2001; Tadic et al., 2007; Dlugos et al., 2009; Unschuld et al., 2009; Lee et al., 2010b; Luo et al., 2010; Rietschel et al., 2010; Aragam et al., 2011; Galeotti and Ghelardini, 2011; Shyn et al., 2011; Dekker et al., 2012; Sakaida et al., 2013; Ray et al., 2014; Gatt et al., 2015; Gray et al., 2015; Ogawa and Kunugi, 2015; Skoog et al., 2015; Wagner et al., 2015; Yoshimasu et al., 2015; De Vry et al., 2016; Liu et al., 2016; Stachowicz et al., 2016) for bioinformatics analysis.

mRNA and ensure that the message is made into protein only when needed (Fernandez-Moya et al., 2014). To date, little work has been done to determine the role of RNA binding factors in antidepression efficacy. Considering the importance of new synaptic connections, RBPs that regulate synapse structure and function are likely to be important for long-lasting antidepressant effects with NMDAR antagonism. Here we will highlight some of the RBPs that are critical for new synapse formation and synaptic scaling.

(i) *NMDAR-mediated RNA granule formation*: mRNA-silencing foci are RNA granules that contain silenced mRNAs and RBPs that repress translation. Distinct mRNA-silencing foci are formed and dispersed with different forms of synaptic plasticity (Thomas et al., 2014). One particular foci that forms with NMDAR stimulation is the synaptic XRN1 bodies (SX-bodies). While XRN1 is an RBP that is best known for its role in mRNA decay, SX-bodies lack decapping activity and therefore serve to sequester mRNAs (Luchelli et al., 2015). Since these bodies increase upon NMDAR stimulation, it is interesting to ask, what happens to SX-bodies with NMDAR antagonism? Moreover, are the sequestered mRNAs needed for the induction of rapid antidepressant efficacy?

(ii) *mRNA trafficking*: A network of RBPs have been implicated in dendritic and synapse morphology. Remarkably, there is little functional redundancy among them. A single knockdown or a mutation in any one RBP disrupts translational regulation of its target mRNAs resulting in dendritic/spine defects. Some of the most compelling data for the importance of RBPs in synapse formation is the knockdown of the RBPs that affect mRNA trafficking (Doyle and Kiebler, 2011). The Staufen proteins are best known for their role in mRNA trafficking into the

dendrites and are required for synaptic plasticity (Tang et al., 2001). Knockdown of Stau2 results in 40% fewer mRNAs in dendrites that are retained in the soma. Moreover, Stau2-knockdown neurons show a significant reduction in synapse number (Goetze et al., 2006). Many depression studies have examined total mRNA expression; however synaptic changes in mRNA levels may be obscured. To the best of our knowledge, none have examined synaptic mRNA levels. Future studies regarding RBP-mediated mRNA trafficking deficits with MDD and with rapid antidepressant treatment are needed.

(iii) *Dendrite and Spine Morphology*: Neurodegenerative diseases with synaptic deficiencies may serve as a clue as to which RBPs may be required for new synapse formation with rapid antidepressants. For example, mutations in two nuclear proteins TAR DNA-binding protein 43 (TDP-43) and fused in sarcoma (FUS) are associated with frontotemporal dementia and amyotrophic lateral sclerosis (ALS). Newly discovered roles for TDP-43 and FUS in dendritic branching and synapse formation, respectively, are emerging (Lu et al., 2009; Sephton et al., 2011). For example, FUS accumulates in dendritic spines to regulate morphology upon mGluR5 stimulation (Fujii and Takumi, 2005). Mutations in FUS result in reduced dendritic branching and spine maturation (Sephton et al., 2014). One might predict that even modest changes in the expression or subcellular localization of TDP-43 and FUS may contribute to the observed synaptic dysfunction, synapse loss, and neuronal atrophy in MDD.

(iv) *Synaptic Scaling*: Two RBPs have been implicated in homeostatic synaptic scaling, the retinoic acid receptor (RAR) α and fragile X mental retardation protein

(FMRP). Both RBPs repress translation of their target mRNAs. During synaptic scaling the increase in GluA1 synthesis arises from the upstream intracellular synthesis of retinoic acid (RA), which in turn, causes dissociation of RARα from GluA1 mRNA, allowing it to be translated (Aoto et al., 2008). Interestingly, Soden and Chen demonstrated that AMPAR-mediated synaptic scaling is absent in CA1 hippocampal neurons of Fmr1 knockout mouse. Moreover, RA-induced mRNA translation of GluA1, GluA2, and eEF2 observed in hippocampal neurons isolated from wildtype mice does not occur in Fmr1 knockout neurons (Soden and Chen, 2010). How FMRP regulates RA-induced synaptic scaling is still an open question. Furthermore, it is unknown whether FMRP- and RA-mediated mRNA translation is required for the rapid antidepressant efficacy induced by NMDAR antagonists.

The RBPs discussed here is a representative sample of RBPs that may be essential to restore proper communication between neurons that will relieve depressive symptoms. Many more are likely to be required. Future work examining the roles of RBPs with NMDAR antagonists is therefore crucial. Moreover, targeting specific RBPs may be critical in revealing more specific therapies for MDD.

CONCLUSION

Rapid antidepressants activate local homeostatic processes. The efficacy of these drugs relies on their ability to engage homeostatic mechanisms that trigger protein synthesis pathways. Interestingly the antidepressant efficacy remains long after the drug has left the patient's system (Abdallah et al., 2015). The acute dose of NMDAR antagonists engages numerous synaptic mechanisms that produce a sustained and lasting increase in

synaptic connections that require protein synthesis. Perhaps these new synapses can reestablish an activity state in the brain that remediates a depressed system. Utilizing native mechanisms of homeostasis, rapid antidepressants offer a novel treatment that may have many applications outside of MDD.

Appendix B: FMRP Regulates an Ethanol-dependent Shift in GABABR Function and Expression with Rapid Antidepressant Properties[11,12]

INTRODUCTION

The presence of major depression increases the risk of alcohol use disorders (AUD) by ~2 fold (and vice versa) (Boden and Fergusson, 2011). The self-medication hypothesis suggests AUDs may develop when the initial antidepressant actions of alcohol are shifted to depressant allostatic states with chronic abuse (Koob and Le Moal, 2001). The molecular mechanism underlying the initial anti-depressant effects of alcohol is unknown.

A major advance in understanding and treating depression is the recognition that NMDA receptor (NMDAR) antagonists act as rapid and effective antidepressant drugs (Skolnick et al., 2009). A single injection of an NMDAR antagonist or "rapid antidepressant" is effective within 2 hours and has sustained antidepressant efficacy for 2 weeks (Abdallah et al., 2015). These long-lasting properties depend on the activity of mammalian target of rapamycin (mTOR), a serine/threonine kinase essential for mRNA translation (Li et al., 2010; Workman et al., 2013). Recently, we demonstrated that activation of mTOR-dependent protein synthesis by NMDA antagonists requires a shift in $GABA_BR$ signaling from opening potassium channels to facilitating an increase in dendritic calcium (Workman et al., 2013; Workman et al., 2015). Interestingly, both acute ethanol and rapid antidepressants block NMDARs (Lovinger et al., 1989) (Abdallah et al., 2015). In light of these data, we propose that ethanol has lasting antidepressant efficacy,

[11] Originally published in Wolfe, S.A., Workman, E.R., Heaney, C., Niere, F., Namjoshi, S., Cacheaux, L.P., Farris, S.P., Drew, M.R., Zemelman, B.V., Harris, R.A., Raab-Graham, K.F. (2016). FMRP Regulates an Ethanol-dependent Shift in GABABR function with Rapid Antidepressant Properties. Nat. Commun. 7: 12867. doi: 10.1038/ncomms12867. I performed FMRP immunoprecipitations and qPCR for figure B.3A.
[12] Supplemental figures available online.

shares the same downstream molecular signaling events as rapid antidepressants, and requires *de novo* protein synthesis (Figure B.S1).

Studies suggest that antidepressant efficacy requires two phases— an induction phase and a sustained phase (Nosyreva et al., 2013). Notably, $GABA_BR$-mediated, mTORC1-dependent protein synthesis is required for the long-lasting sustained phase of rapid antidepressants. Our previous work indicates that both new protein synthesis and an increase in protein stability are required for the $GABA_BR$ shift in function necessary to increase mTORC1 activity (Workman et al., 2015). However, the mechanism that initiates such dynamic changes in protein expression by rapid antidepressants remains unclear.

FMRP is an RNA binding protein that has been characterized as a repressor of mRNA translation. Some forms of synaptic activity trigger FMRP to release its targets, allowing them to be translated(Huber et al., 2002). Moreover, degradation and new protein synthesis of FMRP creates a window for the translation of specific mRNAs, facilitating long-lasting changes in synaptic function (Hou et al., 2006; Nalavadi et al., 2012). Complete loss of FMRP results in Fragile X syndrome (FXS), the single most common genetic cause of autism (Bassell and Warren, 2008). Moreover, reduced levels of FMRP, caused by a pre-mutation, lead to a higher incidence of tremors, ataxia, memory loss, and neuronal neuropathy in older men (Hagerman and Hagerman, 2013). These findings argue that precise levels of FMRP protein and its target mRNAs are required for normal neuronal function.

Drugs of abuse promote profound changes in gene expression, mRNA translation rates, and synaptic protein composition (Ron and Messing, 2013; Placzek et al., 2016). Some studies suggest that drugs and alcohol highjack the molecular mechanisms that underlie synaptic plasticity (Luscher and Malenka, 2011; Lovinger and Roberto, 2013). In agreement with this premise, FMRP has been implicated in cocaine addiction (Smith et al.,

2014). However, little is known about the mRNA targets and the signaling mechanisms involved. Here we describe a critical role for FMRP in mediating $GABA_BR$ synthesis and plasticity following acute ethanol exposure, a mechanism required for antidepressant efficacy.

EXPERIMENTAL PROCEDURES

Cell culture

Primary hippocampal neurons were prepared as previously described by Niere et al., 2012 (Niere et al., 2012). Briefly, hippocampi were extracted from postnatal day 1-3 Sprague-Dawley rat pups, wildtype C57BL/6 mouse pups, or fmr1-knockout (Fmr1 KO) mouse pups on a C57BL/6 background. The tissue was dissociated and plated in neurobasal A medium supplemented with B27, glutamine, and 1% fetal bovine serum. Cultures were plated at a density of ~100,000 cell/12mm on glass coverslips that had been coated overnight with 50 μg/ml poly-D-lysine and 25 μg/ml laminin in borate buffer. Cultures were fed after 1 day in vitro (DIV), and media was replaced approximately once a week with either fresh rat culture media (neurobasal A supplemented with B27, glutamine, and 3 μM AraC) or fresh mouse culture media (glial-conditioned media with 3 μM AraC) until cultures were used at DIV 14-21.

In vitro pharmacology

All pharmacological reagents were added to the incubation media before calcium imaging or immunostaining. Ethanol was added directly to HEPES-based artificial cerebral spinal fluid (ACSF in mM: 100 NaCl, 10 HEPES (pH 7.4), 3 KCl, 2 CaCl2, 1 $MgCl_2$, 10 glucose) that was adjusted to match the osmolarity of cell culture media for live-imaging,

or conditioned cell culture media described above for immunostaining. For pre-treated cultured hippocampal neurons were pre-treated with ethanol (30 mM), cycloheximide (50 μM), or the NMDAR antagonist Ro 25-6981 (10 μM). For $GABA_BR$ activation, neurons were treated with (R)-baclofen (50 μM).

In vitro ethanol treatment

Primary hippocampal neurons were treated in ethanol vapor chambers according to a method adapted from Chandler et al., 1993 (Chandler et al., 1993). Ethanol vapor chambers were prepared by placing a reservoir of 31.5 mM ethanol (105% of the desired ethanol concentration, i.e., 30 mM) in a plastic container with 24-well culture plates containing neuronal cultures in which 30 mM ethanol was added to the culture media. Chambers were filled with 95% O2/5% CO_2 and cultures were incubated for 2 hours at 37°C. Cultures treated with vehicle (H_2O) were incubated in the same manner but in the absence of ethanol. All cultures were treated at 14-21 DIV. Following treatment, cultures were immediately fixed or live imaged.

In vivo ethanol treatment

Male C57BL/6 mice (Charles Rivers) or fmr1 knockout (Fmr1 KO) mice on a C57BL/6 background (at least 7 weeks old) were given intraperitoneal (i.p.) injections of either 200 μl saline or 2.5 g/kg ethanol (in a volume of 200 μl saline) (Blednov et al., 2013). The hippocampi were isolated 30 minutes post-injection and flash frozen. All experiments were carried out in accordance with the National Institutes of Health's Guide for the Care and Use of Laboratory Animals and approved by the UT-Austin Institutional Animal Care and Use Committee (IACUC).

Immunofluorescence

Primary neuronal cultures on glass coverslips were fixed with 4% paraformaldehyde (PFA) at room temperature (RT) for 20 minutes, washed 3 times with phosphate buffered saline (PBS), and permeabilized in 0.25% Triton X-100 in PBS for 5 minutes. For FMRP staining, neurons were fixed and permeabilized in 100% methanol at -20°C for 10 minutes. Neurons were washed 3 times in PBS and then blocked (10% normal goat serum in PBS) for 30 minutes at RT. Primary antibodies were incubated in blocking buffer at 4°C overnight. Neurons were washed 3 times for 10 minutes with PBS, and then incubated in secondary antibody in blocking buffer for 1 hour at RT, and washed 3 times for 10 minutes with PBS before mounting in Fluoromount with DAPI (SouthernBiotech, 0100-20). Surface staining was performed similarly. Neurons were first fixed in 4% PFA for 10 minutes on ice, washed 3 times in PBS, blocked with 3% normal goat serum, and then incubated in primary surface antibody in 3% blocking buffer overnight at 4oC. Following primary surface antibody incubation, neurons were washed 6 times for 10 minutes each in PBS, then permeabilized with 0.25% Triton X-100 in PBS for 5 minutes, followed by 3 washes for 10 minutes each in PBS, and again incubated in primary total antibody in 3% blocking buffer overnight at 4°C. Neurons were washed 3 times for 10 minutes each with PBS, and then incubated in secondary antibody in 3% blocking buffer for 1 hour at RT, and finally washed 4 times for 10 minutes with PBS before mounting in Fluoromount with DAPI to slides. The primary antibodies used were: Total $GABA_BR1$ (1/50 dilution; Santa Cruz, sc-14006), Surface $GABA_BR1$ (1/200 dilution; Abcam, ab55051), $GABA_BR2$ (1/100; Neuromab 75-124), FMRP (1/500 dilution; Abcam ab17722), MAP2 (1/2000 dilution; Abcam ab5392), GFP (1/1,000 dilution; Aves, GFP-

1020). Secondary antibodies included: Alexa488, 555, and 647 developed in goat (1/500 dilution; Life Technologies, A-11039, A-11017, A-31621, A-21430, A-21449, A-21240, A-21237).

Adeno-associated viral vectors

The FMRP and tdTomato coding sequences were cloned into separate adeno-associated viral (AAV) vectors containing a mouse synapsin promoter, a woodchuck post-transcriptional regulatory element (WPRE), and SV40 poly-adenylation sequence between flanking AAV2 inverted terminal repeats (Cajigas et al., 2012). rAAVs were assembled using a modified helper-free system (Stratagene) as serotype 2/1 (rep/cap genes) viruses, and harvested and purified over sequential cesium chloride gradients as previously described (Darnell et al., 2011). Viral titers were greater than 1x109 infectious particles per microliter. For FMRP and tdTomato co-infections, rAAV:mSYN-FMRP and rAAV:mSYN-tdTomato were mixed at a ratio of 4:1. One microliter of the resulting rAAV mix was used per coverslip of primary cultured neurons. Imaging was performed ~1 week after infection.

Live calcium imaging

Dissociated hippocampal cultures were prepared from wildtype (WT) and Fmr1 KO mice as described (Niere et al., 2012). Neurons at 14-21 DIV were used for live calcium imaging. Before imaging, cells were incubated in ACSF with Oregon Green 488 BAPTA-1 AM (OGB, 200 μM; 30 min; 37°C; ThermoFisher) as described (Workman et al., 2013). After OGB incubation, cells were transferred to fresh ACSF (37°C) for imaging (1 frame/20 s). Baseline calcium signal was imaged (1 min), after which (R)-baclofen (50

μM, Tocris) or vehicle (H_2O) was added. For ethanol-treated cells, the neurons were incubated with OGB and imaged in ACSF containing ethanol (30 mM). Neurons were imaged for 800 s at RT. Quantification of the calcium signal was performed using Metamorph (Molecular Devices) as described (Workman et al., 2013). Briefly, dendritic regions of interest (ROI) that were at least 5 μm from the soma were analyzed. The mean intensity values for each ROI at each time were averaged as baseline (F0). The ROI intensity values obtained at each time point after the addition of baclofen or vehicle were averaged (F). The equation, $\Delta F/F = ((F - F0)/F0)$, was used to measure the change in signal and data were plotted as a percentage of the baseline.

BONCAT-PLA

BONCAT-PLA was performed using Click-it Metabolic Labeling AHA, Biotin-Alkyne, and Click-iT Reaction buffer kit (Life Technologies). Proximity ligation assay (PLA) was performed using Duolink kit (Duolink, Sigma) (Ron and Messing, 2013). Briefly, primary hippocampal neuronal cultures were incubated in a methionine-free artificial cerebral spinal fluid (ACSF) media for 30 minutes. AHA was then added to the media just before neurons were treated with ethanol for 2 hours as previously described. Neurons were fixed in 4% PFA for 15 minutes, washed 2 times for 5 minutes with 3% Bovine Serum Albumin (BSA) in PBS, followed by permeabilization with 0.25% Triton X-100 in PBS for 15 minutes, and washed as before. Neurons were incubated for 30 minutes at RT in Cell Buffer Additive/Click-it Cocktail according to manufacturer directions. Neurons were washed as before and then blocked and incubated with primary antibody as previously described. Next, neurons were incubated in the appropriate PLA probes diluted in blocking buffer and secondary antibody at 37°C for 1 hour. Neurons were

washed in RT Buffer A 2 times for 5 minutes, and incubated in ligation solution at 37°C for 30 minutes, and washed again in Buffer A. Neurons were incubated in amplification solution at 37°C for 2-3 hours, followed by washing in RT Buffer B 2 times for 10 minutes and 1% buffer B for 1 minute. Lastly, neurons were mounted to slides in Duolink mounting media for imaging. Primary and secondary antibodies included: $GABA_BR1$ (1/50 dilution; Santa Cruz, sc-14006), $GABA_BR2$ (1/100; Neuromab 75-124), MAP2 (1/2000 dilution; Abcam, ab5392), biotin/α-rabbit (1/500; Sigma, SAB3700857), Alexa488 (1/500; Life Technologies, A-11039). PLA probes used: Rabbit Minus (1/5; Duolink, 82005), Mouse Plus (1/5; Duolink, 82001).

Microscopy and analysis

Images were acquired with a Leica SP5 confocal microscope under a 63X oil immersion lens for fixed tissue or a 63X water immersion lens for live imaging. Max projected images were used for immunostaining analysis from 10-µm Z-stacks of 1024 x 1024 pixels obtained using a 400-Hz scan rate. For each experiment, all microscope settings were held constant. Fixed images were analyzed using NIH imaging software ImageJ, and live imaging quantification was performed with Metamorph (Molecular Devices, Sunnyvale, CA). Relative signal intensity was quantified from dendritic traces. Background signal was subtracted from the traced ROI that lies close to the dendrite (but that is void of all processes) with a custom R script. Dendrites were averaged every 10 µm.

Western blot analysis

Protein was isolated from hippocampal synaptoneurosomes (SN) prepared from mice treated with ethanol or vehicle as previously described. SNs were prepared by

homogenizing hippocampal tissue in homogenization buffer (20 mM HEPES pH 7.4, 5 μM EDTA pH 8.0, and protease inhibitor cocktail). Homogenate was filtered through a 100-μm nylon filter followed by a 5-μm filter, and centrifuged at 14000 x g for 20 minutes at 4°C. The pellet was resuspended in RIPA buffer (150 mM NaCl; 10 mM Tris pH 7.4, 0.1% SDS, 1% Triton X-100, 1% deoxycholate, 5 mM EDTA, and protease inhibitor cocktail) and centrifuged for 10 minutes. The supernatant was then isolated for Western blot analysis. Protein was separated on a 4-20% gradient SDS-polyacrylamide gel. The gel was then transferred to a nitrocellulose membrane, blocked in 5% non-fat dry milk in tris-buffered saline and 0.1% tween-20 (TBST) for 1 hour, and incubated with primary antibody in blocking buffer overnight at 4°C. The blot was washed in TBST 3 times for 10 minutes each, incubated in secondary antibody for 1 hour, and washed as before. Blots were imaged using a LICOR Odyssey imaging system, and ImageJ was used for densitometry analysis. Representative images are pseudocolored with black (lowest intensity) to red (highest intensity) using LICOR Image Studio software. Primary antibodies used consisted of: $GABA_BR1$ (1/400 dilution; Santa Cruz, sc-14006), $GABA_BR2$ (1/800; Neuromab 75-124), alpha-Tubulin (1/2000 dilution; Sigma, T6074). Secondary antibodies included: anti-mouse-IR-Dye 800 (1/5000 dilution excluding tubulin at 1/10,000 dilution; LICOR, 96-32210) and anti-rabbit Alexa680 (1/5000 dilution; Invitrogen, A-21084).

RNA Immunoprecipitations

Cortices from 6-week old C57BL/6 mice were harvested and flash frozen on dry ice. Tissue was homogenized and lysed with a cordless pestle motor and disposable pellet mixers (VWR) in polysome lysis buffer (10 mM HEPES pH 7.0, 100 mM KCl, 25 mM

EDTA, 5 mM MgCl2, 1mM DTT, 0.5% NP-40) in a 1:1 tissue-buffer ratio. RNaseOUT (Thermo) and protease/phosphatase inhibitors (Halt Protease and Phosphatase Cocktail, Pierce Biotechnology) were freshly added to samples. Samples were rotated for 10 minutes at 4°C to induce swelling and then flash frozen on dry ice. Samples were thawed by holding between fingers at RT to lyse and nuclei were pelleted at 3000 x g for 10 minutes. Lysates obtained above were pre-cleared by adding 50 µl of washed magnetic bead slurry (Protein A Dynabeads, Thermo) and rotating for 30 minutes at 4°C. To bind the antibody to the beads, 50 µl of magnetic beads slurry was washed and then resuspended in 8 volumes of NT-2 buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM MgCl2, 1mM DTT, 0.05% NP-40 with RNaseOUT/protease & phosphatase inhibitors added fresh) + 5% BSA. 10 µg of either FMRP (Abcam, ab17722) or IgG (Santa Cruz Biotechnologies, sc-2027) antibodies were added to the beads and rotated for 10 minutes at RT. Antibody-bound beads were washed 4 times with ice-cold NT-2 buffer. For the immunoprecipitation, 4.5 mg of protein from pre-cleared lysates was added to an RNase-free microcentrifuge tube containing the antibody-bound beads. Input collected at this step for downstream analysis was either 1% of the final pre-cleared lysate volume in the immunoprecipitation reaction (for immunoblotting) or 10% of the final pre-cleared lysate volume (to normalize in qPCR). The antibody-bead-lysate mixture was then diluted at a ratio of 1:5 with NET-2 buffer (20 mM EDTA pH 8.0, and 1 mM DTT in NT-2 buffer; RNaseOUT and protease/phosphatase inhibitors added fresh) and rotated for 1 hour at room temperature. Beads were quickly washed 6 times in ice-cold NT-2 buffer and immediately resuspended in 350 µl TRI Reagent Solution (Ambion) for 10 minutes at RT. Beads were pelleted and the supernatant was removed and resuspended in 350 µl of absolute ethanol. RNA was extracted by applying ethanol-resuspended samples to spin column from the Direct-zol RNA MiniPrep

Kit (Zymogen) according to manufacturer's instructions. Eluted RNA (25 µl) was DNase treated using the TURBO DNA-free kit (Thermo).

cDNA Synthesis and Quantitative Real-Time PCR (qRT-PCR)

DNase-treated RNA samples were reverse-transcribed to cDNA using the iScript cDNA Synthesis Kit (Bio-Rad) in a 20 µl volume according to manufacturer's instructions. qRT-PCR was performed in 20 µl reaction volume using the iQ SYBR Green Supermix (Bio-Rad) and primers for $GABA_BR1$, $GABA_BR2$, CaMKIIα, and Cacna2δ2 (GeneCopoeia). qRT-PCR was run with the following protocol: 95°C for 10:00, 40 cycles of 95°C for 0:15 followed by 60°C for 1:00, 95°C for 1:00, and 55°C for 1:00. Relative fold-enrichment was determined by the equation $\Delta\Delta Ct = 2^{-(Ct\ FMRP\ RNA\text{-}IP - Ct\ IgG\ RNA\text{-}IP) - (Ct\ FMRP\ input - Ct\ IgG\ input)}$.

Forced Swim Test

Mice were tested in the forced swim test 24 hours post-injection as described previously 35-37. Mice were individually placed into a cylinder containing 3 L of water (25°C) for 6 minutes. Each session was video recorded and the last 4 minutes of the sessions were later scored blindly for immobility. Animals were scored for escape-directed behaviors. The water was replaced between animals. Experiments were repeated by 3 independent experimenters. Data was normalized by experimenter.

Open Field

Twenty-four hours after animals were injected i.p. with either saline or ethanol, they were studied in the open field test. Mice were individually placed in a 40 cm x 40 cm

x 35 cm arena with opaque walls. Each test session lasted 30 minutes under 85 lux illumination. Sessions were video recorded and analyzed via ANY-Maze (Stoelting, Wood Dale, IL). Mice were considered to be in the center of the maze if they entered an 18.5 cm x 18.5 cm area in the center of the apparatus. Mice were returned to their home cage at the end of the test session, and the arena was wiped down with 70% ethanol before the start of each run.

Splash Test

Two and a half hours after open field testing, animals underwent the splash test. Cage mates were moved from their home cage to a holding cage, and each animal was individually tested in its home cage. Two hundred microliters of 10% sucrose was applied to the dorsal fur of the mouse. Mice were monitored and video recorded for 5 minutes and then moved to a different holding cage. Videos were later scored blindly for latency to initiate grooming and for total time spent grooming. Grooming behavior included licking, grooming with forepaws, and scratching.

Statistical analysis

Prism software (GraphPad) was used for all statistical analyses. Statistical comparisons were made using one-way ANOVA, two-way ANOVA, Student's t-test, or χ^2 test with Yates. Outliers were determined using Grubbs' test (alpha=0.05). All data are expressed as mean ± Standard Error of the Mean (SEM). Significance is represented by: n.s., not significant, *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$ within-genotype comparisons, and # indicates significance between genotypes for Figure B.S6 (#$p \leq 0.05$).

RESULTS

Antidepressant and anxiolytic effects of ethanol on behavior

To determine if acute alcohol has antidepressant properties, as predicted by the self-medication hypothesis, we first assessed the efficacy of alcohol on antidepressant- and anxiolytic-like effects on behavior. The forced swim test (FST) is a behavioral test of despair (Porsolt et al., 1979). Rodents treated with a single injection of NMDAR antagonists or rapid antidepressants swim longer and thus have reduced times of immobility relative to controls. Notably, these positive effects on behavior last long after the drug has been metabolized (Li et al., 2010; Autry et al., 2011; Workman et al., 2013; Workman et al., 2015). Therefore, we considered the possibility that ethanol, which blocks NMDARs (Lovinger et al., 1989), could also act like an antidepressant at 24 hours, well beyond the intoxication period (Blednov et al., 2013). To test this, C57BL/6 mice were injected with ethanol (2.5 g/kg, i.p.), a concentration that is achieved during self-administration in mice (Rhodes et al., 2005). Twenty-four hours after injection, the immobility time in ethanol-treated mice was reduced by ~15% relative to controls (Figure B.1a), similar to our previous observation in mice that had been exposed to the rapid antidepressant Ro 25-6981 (Workman et al., 2013; Workman et al., 2015). These results demonstrate that acute ethanol elicits an antidepressant effect on behavior similar to that seen with rapid antidepressants (Workman et al., 2015).

As another measure of antidepressant effect of ethanol on behavior, we assessed the grooming behavior of mice using the splash test after ethanol or saline administration. The splash test measures latency to groom and dedicated grooming time as indicators of self-care and motivational behavior (Surget et al., 2008; David et al., 2009). Lack of self-care is often observed in humans with depressive disorder (Association, 2013). We have previously shown that mice receiving a single i.p. injection of the rapid antidepressant Ro

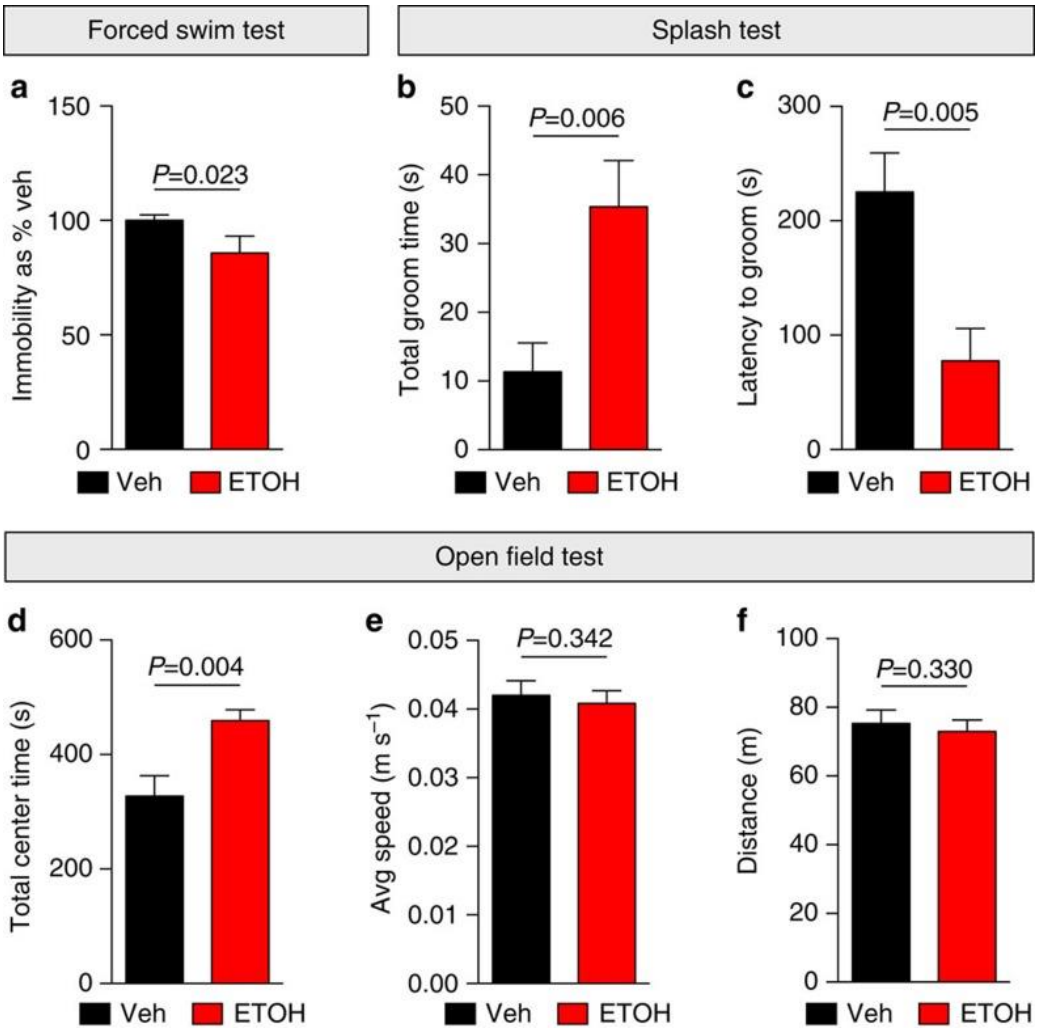


Figure B.1: Ethanol elicits a lasting antidepressant-like effect on behavior

a. C57BL/6 male mice were subjected to the forced swim test 24 h after i.p. injection with vehicle (Veh; saline) or ethanol (ETOH; 2.5 g kg^{-1}). Ethanol treatment reduced immobility, indicating antidepressant efficacy. Veh=100±2.5, *n*=10 mice; ETOH=86±7.4, *n*=6. **b**. **and c.** In the splash test, male C57BL/6 mice groomed longer and took less time to initiate grooming 24 h post ethanol (2.5 g kg^{-1}, i.p.) compared with 24 h post-vehicle (saline, i.p.) treatment. Total groom time: Veh=11.34±4.23 s, *n*=6; ETOH=35.37±6.72 s, *n*=5. Latency to groom: Veh=225.2±34.13 s, *n*=6; ETOH=77.55±28.44 s, *n*=5. **d.**–**f.** Total centre time, speed and distance were measured in the open field test 24 h post injection. Ethanol-treated (2.5 g kg^{-1}, i.p.) mice spent more time in the centre, while speed and distance were unaffected compared with vehicle-treated (saline, i.p.) mice, indicating an ethanol-induced anxiolytic effect without altering mobility. Total centre time: Veh=327.5±35.62 s, *n*=6; ETOH=459.2±19.13 s, *n*=6. Average speed: Veh=0.042±0.002 m s^{-1}, *n*=6; ETOH=0.041±0.002 m s^{-1}, *n*=6. Total distance: Veh=75.35±3.92 m, *n*=6; ETOH=73.00±3.35 m, *n*=6. Significance determined by one-tailed *t*-test. Values represent mean±s.e.m.

25-6981 spend more time grooming compared to control mice (Workman et al., 2015). We hypothesized that ethanol would produce similar effects on grooming behavior. Indeed, ethanol-treated mice spent more time grooming and displayed shorter latency to initiate grooming relative to controls (Figure B.1b and c).

Ethanol is a well-known anxiolytic substance (Blednov et al., 2001). However, the anxiolytic effect of a single dose of ethanol 24 hours after administration has not been determined. We subjected ethanol- and saline-injected mice to the open-field test to assess the influence of ethanol on anxiety-like behaviors after 24 hours. Mice that spend more time in the center of the open field are scored to indicate reduced anxiety-like behavior relative to mice that remain close to the perimeter (Treit and Fundytus, 1988). Indeed, mice that received a single dose of ethanol (2.5 g/kg, i.p.) had reduced anxiety-like behavior, spending ~40% more time in the center relative to controls (Figure B.1d). There was no significant difference in total distance traveled or average speed between the groups (Figure B.1e-f). These data suggest that the anxiolytic effects of ethanol last up to 24 hours post-injection.

Acute ethanol increases $GABA_BR2$ and surface $GABA_BRs$

Both ethanol and rapid antidepressants block NMDARs in the hippocampus (Lovinger et al., 1989; 1990). One of the first events triggered by NMDAR antagonism is increased dendritic $GABA_BR2$ protein expression (Workman et al., 2015). $GABA_BRs$ are obligate heteromultimers, consisting of $GABA_BR1$ and R2. $GABA_BR2$ is required for expression of receptors at the surface by masking an endoplasmic reticulum retention sequence on $GABA_BR1$ (Margeta-Mitrovic et al., 2000). Similarly, treatment with a rapid antidepressant leads to (1) increased dendritic expression of $GABA_BR2$ but not $GABA_BR1$

(Workman et al., 2015), and (2) a corresponding increase in surface expression of GABABR1 (Sosanya et al., 2015b; Workman et al., 2015).

To determine if acute ethanol exposure in vivo rapidly increases the levels of $GABA_BR1$ and/or $GABA_BR2$, hippocampal synaptoneurosomes were isolated from mice that had been injected with a single dose of ethanol (2.5 g/kg, i.p.) or saline for Western blot analysis. The hippocampi were collected within the initiation phase (30 minutes post-injection), a phase where molecular changes facilitate increased downstream mTORC1 activity (Raab-Graham et al., 2016). Consistent with rapid antidepressants, acute ethanol injection increased the protein expression of $GABA_BR2$ by ~37% in the hippocampus with no significant change in $GABA_BR1$ (Figure B.2a-c).

To further identify the subcellular localization of ethanol-induced increase in $GABA_BR2$, we examined $GABA_BR$ expression in cultured hippocampal neurons. $GABA_BR1$ and R2 were immunostained and quantified in the dendrites. A concentration of 30mM ethanol was chosen, as it has been shown to reduce NMDAR activity in hippocampal neurons and reflects that achieved in vivo following i.p. injection (Lovinger et al., 1989). Acute ethanol exposure (30 mM, 2 hours) increased the dendritic levels of GABABR2 by ~47%, but did not affect $GABA_BR1$ levels (Figure B.2d-g). We did not observe a difference in the diameter of the primary dendrites between vehicle- and ethanol-treated neurons, demonstrating that ethanol does not modify dendritic caliber (Figure B.S2). These *in vivo* and *in vitro* findings establish a role for ethanol in increasing $GABA_BR2$ protein expression.

Since $GABA_BR2$ is required for the surface expression of the heteromultimeric receptor, we predicted that the ethanol-induced elevation in $GABA_BR2$ levels would increase expression of receptors at the surface. We measured the surface expression of dendritic $GABA_BRs$ using an antibody directed against the extracellular domain of

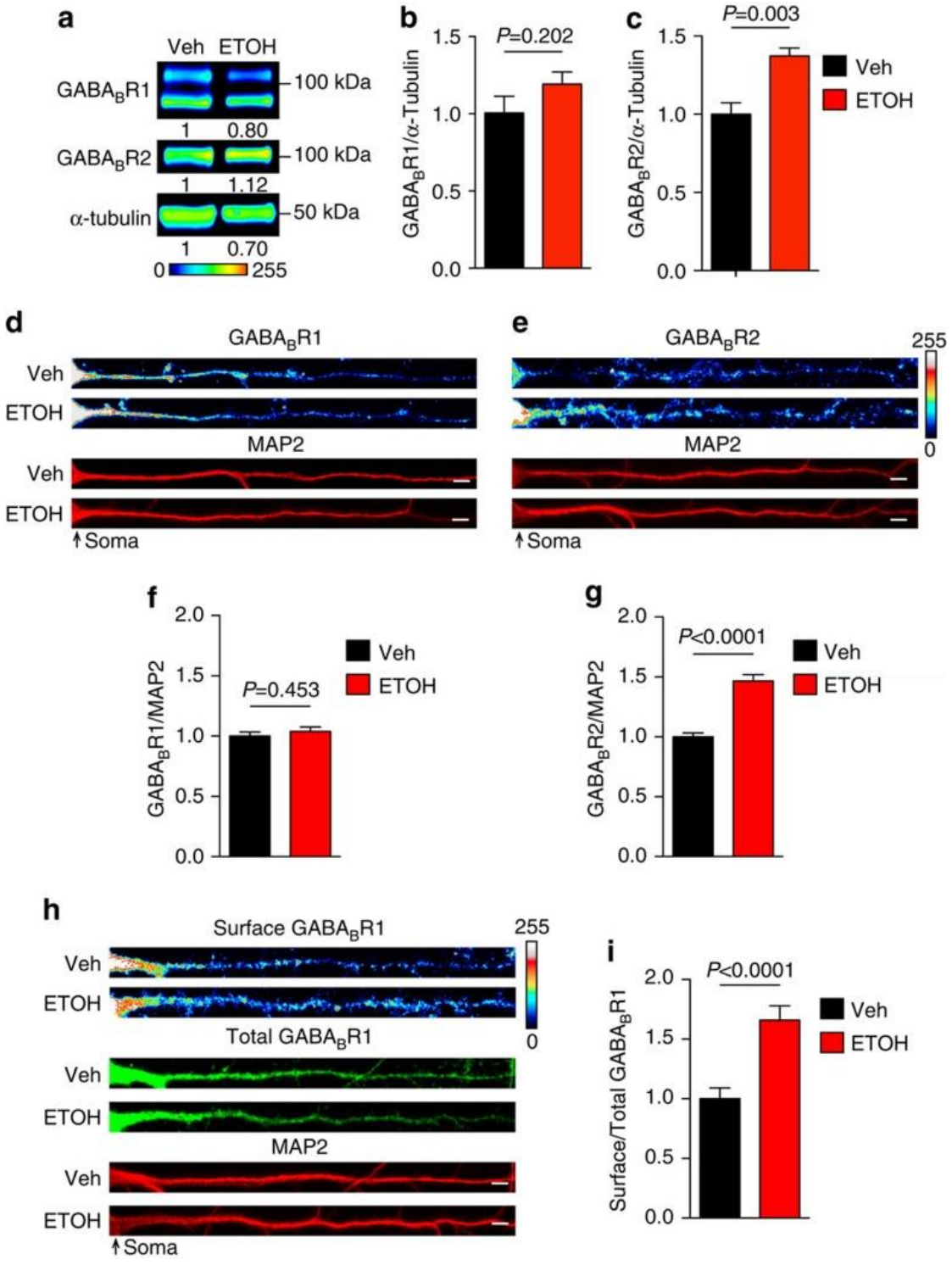


Figure B.2: Acute ethanol increases dendritic GABA$_B$Rs in hippocampus

$GABA_BR1$ in unpermeabilized hippocampal neurons. The surface signal was normalized by the total dendritic $GABA_BR1$ levels after permeabilization6. As predicted, surface expression of $GABA_BRs$ in ethanol-treated neurons was significantly higher (~66% increase) relative to controls (Figure B.2h-i). This ethanol effect was consistent with what we previously observed following rapid antidepressant treatment of cultured hippocampal neurons (Workman et al., 2013). Collectively, these results suggest that ethanol promotes the surface expression of $GABA_BRs$, and this is likely achieved by increasing $GABA_BR2$ protein levels.

FMRP regulates the protein expression of $GABA_BR1$ and $GABA_BR2$

Next, we sought to identify the mechanism by which NMDAR antagonism increases $GABA_BR2$ expression. $GABA_BR2$ mRNA is present in the dendrites of hippocampal neurons (Cajigas et al., 2012), suggesting that this mRNA may be locally

---

a.–**c.** Western blot analyses of $GABA_BR1$ and $GABA_BR2$ in isolated hippocampal synaptoneurosomes from ethanol-treated (ETOH; 2.5 g kg^{-1} i.p.), and vehicle-treated (Veh; saline i.p.) C57BL/6 male mice 30 min post injection. **a**. Pseudocoloured representative western blots to show intensity with normalized optical density for each band indicated below blot (Lookup table, below western blot). No significant change was observed in **b.** $GABA_BR1$, but a significant increase was found in **c** $GABA_BR2$ with ethanol treatment. Western blots were normalized to the loading control, α-Tubulin. $GABA_BR1$: Veh=1.00±0.11; ETOH=1.19±0.08. Experiment was repeated five times. $GABA_BR2$: Veh=1.00±0.07; ETOH=1.37±0.05. **d**. **and e.** Representative immunostaining images of $GABA_BR1$ and $GABA_BR2$ in cultured rat hippocampal dendrites normalized to microtubule associated protein 2 (MAP2) as volume control. There was no change in **f.** $GABA_BR1$ and a significant increase in **g.** $GABA_BR2$ in ethanol treated (30 mM, 2 h) compared with vehicle-treated (H2O, 2 h) dendrites: Total $GABA_BR1$: Veh=1.00±0.03, *n*=46 dendrites; ETOH=1.04±0.04, *n*=51 dendrites. Total $GABA_BR2$: Veh= 1.00±0.03, *n*=46 dendrites; ETOH=1.47±0.05, *n*=51 dendrites. **h**. **and i.** Immunofluorescence shows a significant increase in surface $GABA_BR1$ expression in dendrites of cultured rat hippocampal neurons treated with ethanol (30 mM, 2 h); **i**. Surface expression of $GABA_BR1$ in vehicle-treated (H_2O, 2 h) and ethanol-treated (30 mM, 2 h) dendrites. Veh=1.00±0.09, *n*=43 dendrites; ETOH=1.66±0.12, *n*=47 dendrites. Significance determined by Student's *t*-test. Values represent mean±s.e.m. Scale bars, 5 μm. Uncropped version of western blots, with size markers are available in Figure B.S7a

regulated at the translational level. Thus, we examined RNA-binding factors that may regulate $GABA_BR2$ mRNA expression in dendrites. Notably, both $GABA_BR1$ and $GABA_BR2$ mRNAs are reported targets of FMRP, an RNA binding protein that stalls translational elongation of its targets (Darnell et al., 2011; Graber et al., 2013a).

To test the hypothesis that FMRP regulates $GABA_BR$ mRNA translation, we first verified that (1) $GABA_BR$ mRNAs bind to FMRP, and that (2) the absence of FMRP in knockout mice results in aberrant expression of $GABA_BRs$. Using a specific antibody against FMRP, bound mRNAs were isolated using RNA immunoprecipitation (RNA-IP). $GABA_BR1$ and $GABA_BR2$ binding were assessed by reverse transcription (RT) and quantitative PCR. Indeed, $GABA_BR1$ and $GABA_BR2$ mRNAs were detected in the immunoprecipitate, along with CaMKIIα, a well-known FMRP mRNA target (Figure B.3a-b). The calcium channel accessory subunit Cacnα2δ2 mRNA is not a reported target for FMRP (Darnell et al., 2011) and was used as a negative control. Cacnα2δ2 mRNA was not detected in the FMRP RNA-IP (Figure B.3b). In parallel, we used lysates isolated from brains of mice with a genetic deletion of the Fmr1 gene. We did not observe amplification of any of the mRNAs in Fmr1 KO brains, providing additional evidence for specific binding of FMRP to $GABA_BR1$ and $GABA_BR2$ mRNAs (Figure B.3a-c).

Next, we determined if FMRP regulates $GABA_BR1$ and $GABA_BR2$ protein levels. Genetic deletion of Fmr1 KO leads to the constitutive translation of FMRP target mRNAs and the loss of activity-dependent translation (Bassell and Warren, 2008). Protein levels of $GABA_BR1$ and $GABA_BR2$ were compared in hippocampal synaptoneurosomes from Fmr1 KO and wildtype (WT) mice (Figure B.3c-g). $GABA_BR2$ basal protein levels were elevated by ~53% in Fmr1 KO hippocampi (Figure B.3g). $GABA_BR1$ protein levels also increased, albeit to a lesser extent than $GABA_BR2$ (Figure B.3f). Collectively, these data suggest that FMRP regulates the expression of $GABA_BR1$ and $GABA_BR2$.

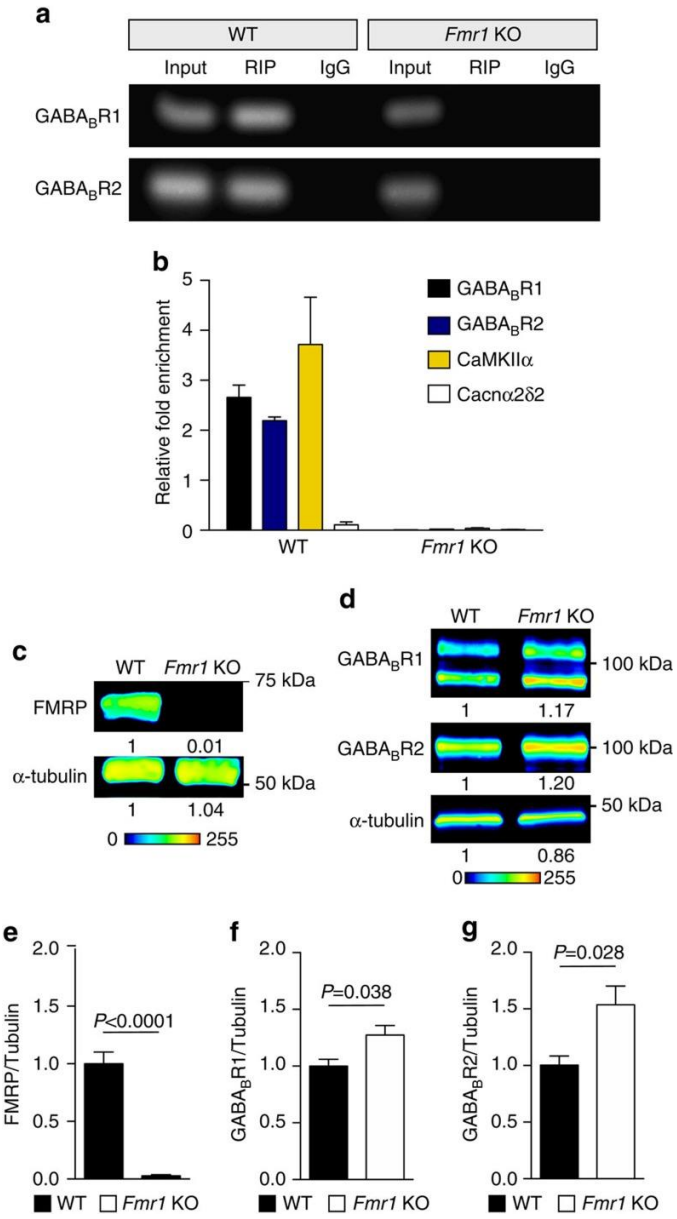


Figure B.3: GABA$_B$R1 and GABA$_B$R2 mRNAs are FMRP targets

a. **and b.** RNA immunoprecipitation (RIP) for FMRP was performed using brains from wild type (WT) and *Fmr1* KO male mice. **a**. Gels showing RT–qPCR amplified product of input sample, FMRP RIP, and IgG control for GABA$_B$R1 and GABA$_B$R2. **b**. Relative fold-enrichment as determined by real-time qPCR relative to input control ($\Delta\Delta Ct=2^{-(Ct\ FMRP\ RIP-Ct\ IgG\ RIP)-(Ct\ FMRP\ input-Ct\ IgG\ input)}$). FMRP binds GABA$_B$R1, GABA$_B$R2, and the positive control CaMKIIα mRNA as detected in the RIP sample by real-time qPCR. Cacnα2δ2 served as a negative control and was not detected above background. WT: GABA$_B$R1=2.66±0.248, n=2; GABA$_B$R2=2.19±0.08, n=2; CaMKII=3.72±0.94, n=2; Cacnα2δ2=0.11±0.6, n=2. *Fmr1* KO: GABA$_B$R1=0.01±0.0002, n=2;

Ethanol and rapid antidepressants reduce dendritic FMRP

As an initial test to determine if FMRP-regulated translation is linked to alcohol exposure, we compared FMRP target mRNAs (Darnell et al., 2011) with mRNAs that are differentially expressed in the hippocampi of alcohol-dependent humans (Zhou et al., 2011). Remarkably, 225 or ~25% of verified FMRP target mRNAs overlap with mRNAs that are altered in alcohol-dependent individuals, suggesting a role for FMRP in aberrant protein levels observed in humans with AUD (Figure B.4a) (Darnell et al., 2011; Zhou et al., 2011). We then determined if exposure to acute ethanol (30 mM, 2 hours) or Ro 25-6981 (10 μM, 2 hours) alters FMRP expression in the dendrites of hippocampal neurons. Using immunofluorescence, we found that ethanol and Ro 25-6981 reduced FMRP levels by ~38% and 45%, respectively (ethanol: Figure B.4b-c; Ro-25-6981: Figure B.S3). These data suggest that ethanol and Ro 25-6981 alter protein expression in an FMRP-dependent manner.

Ethanol-induced synthesis of $GABA_BR2$ requires FMRP

Due to the ethanol-induced decreases in FMRP, we hypothesized that FMRP is required for ethanol-induced expression of $GABA_BRs$. Specifically, if expression of $GABA_BRs$ is constitutive and unregulated in Fmr1 KO mice (Bassell and Warren, 2008),

---

$GABA_BR2$=0.02±0.00006, *n*=2; CaMKII±0.04±0.01, *n*=2; Cacnα2δ2=0.012±0.005, *n*=2. **c.**–**g.** Western blot analysis of hippocampal synaptoneurosomes isolated from C57BL/6 WT and *Fmr1* KO mice on a C57BL/6 background indicates the absence of **e**. FMRP and increased protein expression of **f**. $GABA_BR1$ and **g**. $GABA_BR2$. Representative western blots are pseudocoloured to indicate intensity of bands, and the normalized optical density for each band is indicated below blot (Lookup table, below western blot). Western blots were normalized to the loading control, α-Tubulin. WT: FMRP=1.00±0.10; $GABA_BR1$=1.00±0.06; $GABA_BR2$=1.00±0.08. *Fmr1* KO: FMRP=0.03±0.01; $GABA_BR1$=1.27±0.08; $GABA_BR2$=1.54±0.17. Experiment was repeated three times. Significance determined by Student's *t*-test. Values represent mean±s.e.m. Uncropped versions of qPCR gel, with size markers, are available in Figure B.S7d. Uncropped version of western blots, with size markers are available in Figure B.S7b.

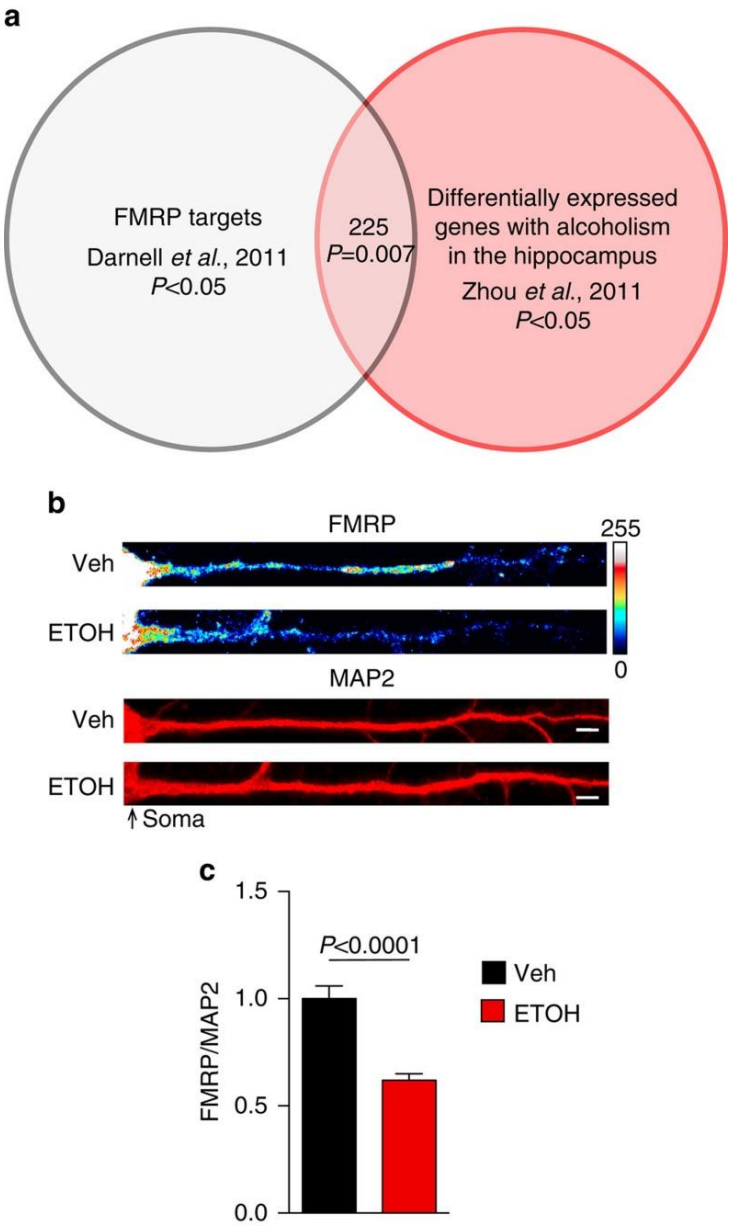


Figure B.4: FMRP and AUD share target mRNAs and ethanol decreases FMRP

a. Venn diagram illustrating the significant overlap between FMRP mRNA targets and differentially expressed genes in the hippocampus of humans with AUDs. Significance determined with χ^2-test. **b**. **and c.** Immunofluorescence images normalized to MAP2 as volume control and quantification summary shows a significant decrease in FMRP expression in dendrites of cultured hippocampal neurons treated with ethanol (ETOH; 30 mM, 2 h) compared with vehicle treated (Veh; H_2O, 2 h). Veh: FMRP =1.00±0.06, n=32 dendrites; ETOH: FMRP=0.62±0.03, n=31 dendrites. Significance determined by Student's t-test. Values represent mean±s.e.m. Scale bar, 5 μm.

then ethanol-induced changes in $GABA_BR$ expression should be absent in Fmr1 KO mice. Hippocampal synaptoneurosomes were isolated from WT and Fmr1 KO mice 30 minutes after i.p. injection of ethanol (2.5g/kg). Western blot analysis indicated that both $GABA_BR1$ and $GABA_BR2$ expression remained constant in vehicle- and ethanol- treated Fmr1 KO mice. As observed in Figure B.2, wildtype WT hippocampal synaptoneurosomes showed an ~23% increase in $GABA_BR2$ but no change in $GABA_BR1$ expression (Figure B.5a-c). These data suggest that ethanol-induced changes in $GABA_BR$ expression are dependent on FMRP translational regulation.

To determine whether protein synthesis is essential for the FMRP-dependent changes in $GABA_BR$ expression, we measured ethanol-induced $GABA_BRs$ in the presence of cycloheximide (CHX), a protein synthesis inhibitor. As demonstrated previously, ethanol did not influence the dendritic expression of $GABA_BR1$; however, co-treatment with cycloheximide increased $GABA_BR1$ expression by ~22%. FMRP deletion did not affect the basal, ethanol-, or cycloheximide-induced dendritic protein expression of $GABA_BR1$ (Figure B.5d-e and h). For $GABA_BR2$, we again saw a significant ~28% increase in dendritic expression with acute ethanol treatment; however, in the presence of cycloheximide the ethanol-induced increase was abolished. Notably, when the same treatments were applied to the Fmr1 KO cultures, no change was observed with ethanol or ethanol+cycloheximide (Figure B.5f-g and i).

We next examined the requirement for protein synthesis and FMRP in ethanol-dependent surface expression of $GABA_BRs$. Using WT and Fmr1 KO hippocampal neurons, we measured ethanol-induced surface expression of $GABA_BR1$ with or without cycloheximide. Co-assembly of $GABA_BR1$ and R2 is required to express $GABA_BR$ heterodimers in the membrane (Margeta-Mitrovic et al., 2000). Thus, we predicted that the ethanol-induced increase in surface $GABA_BRs$ would require FMRP-regulated synthesis

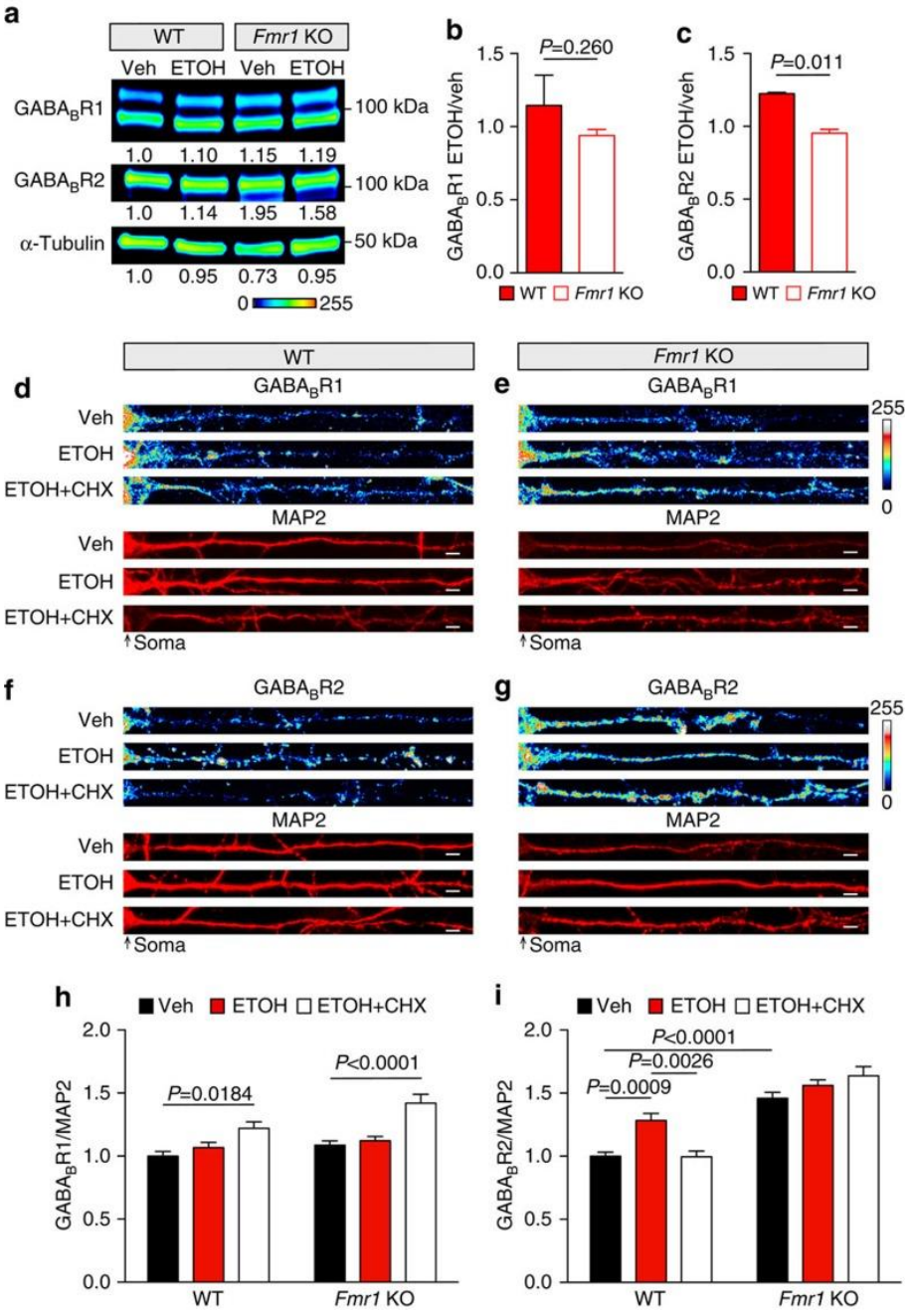


Figure B.5: *Fmr1* KO prevents ethanol-induced altered $GABA_BR$ expression

a.–c. Western blot analysis of $GABA_BR1$ and $GABA_BR2$ in wild type (WT) and *Fmr1* KO C57BL/6 hippocampal synaptoneurosomes after vehicle (Veh; saline i.p., 30 min) or ethanol (ETOH; 2.5 g kg^{-1} i.p., 30 min) treatment. **a**. Pseudocoloured representative western blots showing band intensity, and normalized optical densities to WT–vehicle are reported below each image (lookup table, below western blot). Western blots were normalized to the loading control, α-

of $GABA_BR2$. Again, acute ethanol increased dendritic surface $GABA_BRs$ by ~76%, and this was blocked by cycloheximide. In Fmr1 KO neurons there was no significant ethanol-induced change in surface $GABA_BRs$, but a decrease was observed with ethanol+cycloheximide (Figure B.6a-c). These data suggest that $GABA_BR2$ protein synthesis is required for the increased surface expression of the hetermultimeric receptor with ethanol exposure.

FMRP is reported to repress new protein synthesis. Considering the effects of the protein synthesis inhibitor, these data suggests that the ethanol-mediated reduction in FMRP results in the increase in dendritic $GABA_BR2$ by de novo protein synthesis. To provide more direct evidence, we performed bioorthogonal noncanonical amino acid tagging (BONCAT) in conjunction with proximity ligation assay (PLA-Duolink) (Dieterich et al., 2010). BONCAT+PLA can be used to detect new synthesis of proteins of interest, such as $GABA_BR1$ and $GABA_BR2$. Through click chemistry, noncanonical amino

---

Tubulin. No change was found in **b.** $GABA_BR1$ after ethanol treatment in either genotype as shown by ethanol/vehicle comparison. A significant increase in **c** $GABA_BR2$ expression was observed in WT mice after ethanol, but no change was observed in *Fmr1* KO mice (shown as ethanol/vehicle). WT ETOH/Veh: $GABA_BR1$=1.15±0.21; $GABA_BR2$=1.22±0.01. *Fmr1* KO ETOH/Veh: $GABA_BR1$=0.94±0.04; $GABA_BR2$=0.95±0.03. Experiment was repeated three times. Significance determined by Student's *t*-test. Values represent mean±s.e.m. Representative immunofluorescent images **d.**–**g.** and quantification summaries **h. and i.** of dendritic expression of $GABA_BR1$ and $GABA_BR2$ from WT and *Fmr1* KO primary mouse hippocampal cultures normalized to MAP2. **H**. $GABA_BR1$ expression was not changed in either genotype after 2-h treatment with vehicle (Veh; H_2O), ethanol (ETOH; 30 mM), or ethanol and cycloheximide (30 mM ETOH+50 μM CHX). WT $GABA_BR1$: Veh=1.00±0.04, *n*=44 dendrites; ETOH=1.07±0.04, *n*=29 dendrites; ETOH+CHX=1.22±0.05, *n*=34 dendrites. *Fmr1* KO $GABA_BR1$: Veh=1.09±0.03, *n*=72 dendrites; ETOH=1.12±0.03, *n*=41 dendrites; ETOH+CHX=1.42±0.07, *n*=43 dendrites. **i**. $GABA_BR2$ expression in WT neurons increased after ethanol (ETOH; 30 mM, 2 h) compared with vehicle (Veh; H_2O, 2 h) treatment, and was rescued with co-treatment of cycloheximide (CHX; 50 μM, 2 h). $GABA_BR2$ expression in *Fmr1* KO dendrites was not significantly altered between neurons treated with Veh, ETOH, or ETOH+CHX. WT $GABA_BR2$: Veh=1.00±0.03, *n*=41 dendrites; ETOH=1.28±0.06, *n*=40 dendrites; ETOH+CHX=0.99±0.05, *n*=33 dendrites. *Fmr1* KO $GABA_BR2$: Veh=1.46±0.05, *n*=73 dendrites; ETOH=1.56±0.04, *n*=45 dendrites; ETOH+CHX=1.63±0.07, *n*=36 dendrites. Significance determined by two-way analysis of variance with Tukey's post test. Value represent mean±s.e.m. Scale bars, 5 μm. Uncropped version of western blots, with size markers are available in Figure B.S7c.

acids that are incorporated during mRNA translation are biotinylated. PLA, on the other hand, generates a fluorescent signal when two antibodies are within 30-40 nm of each other (i.e., anti-$GABA_BR1$ or $GABA_BR2$ and anti-biotin). By combining these methods, we determined that ethanol treatment increases new protein synthesis of $GABA_BR2$ by ~40%, but does not alter $GABA_BR1$ synthesis, similar to the $GABA_BR$ changes induced by rapid antidepressants (Workman et al., 2015). In Fmr1 KO dendrites, basal levels of $GABA_BR2$s increased by ~48%, while a significant decrease was observed in $GABA_BR1$ levels. Additionally, ethanol-induced translation of $GABA_BR2$ was lost in Fmr1 KO dendrites (Figure B.6d-i). These data provide additional evidence that the ethanol-induced increase in $GABA_BR2$ expression is due to new protein synthesis that requires the release of translational repression by FMRP.

Ethanol-induced $GABA_BR$ plasticity requires FMRP

We previously demonstrated that rapid antidepressants shift $GABA_BR$ signaling from opening potassium channels to increasing dendritic calcium (Workman et al., 2013). To determine if ethanol (30 mM, 2 hours) causes the same plasticity in $GABA_BR$ signaling, we performed fluorescence calcium imaging in cultured WT and Fmr1 KO hippocampal neurons. A transient rise or fall in calcium in dendritic compartments can be detected using a fluorescent indicator that exhibits changes in fluorescent properties depending on the amount of bound calcium (Hendel et al., 2008). We used baclofen, a $GABA_BR$ agonist, to activate $GABA_BR$s in the presence or absence of ethanol. After establishing a baseline measurement, baclofen reduced dendritic calcium fluorescence in saline-treated WT neurons by ~11%, a characteristic signature of $GABA_BR$ signaling increasing outward potassium conductance (Workman et al., 2015). However, in ethanol-treated WT neurons, baclofen induced distinct calcium waves and an overall averaged increase in calcium signal

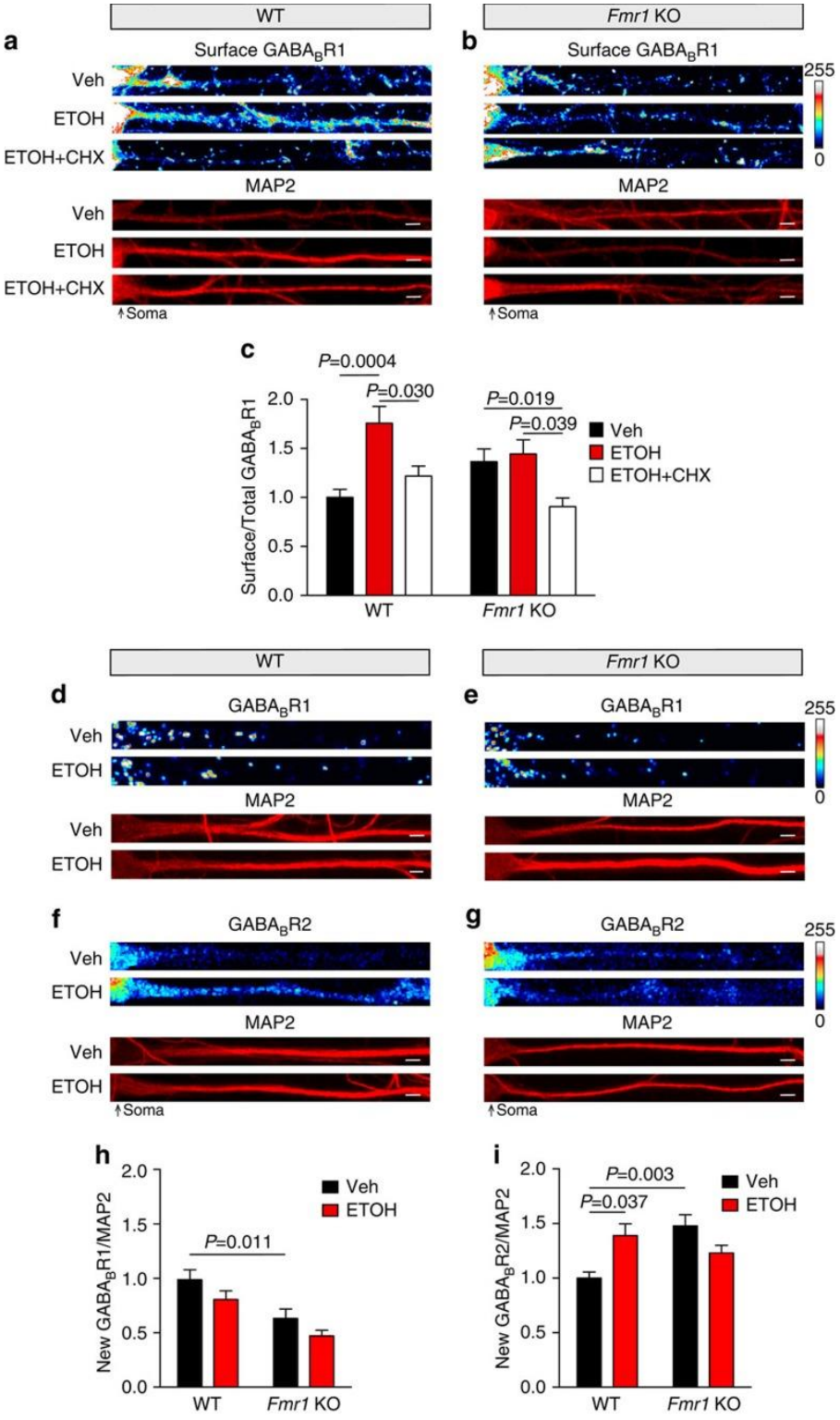


Figure B.6: New $GABA_BR2$ protein and surface expression requires FMRP

of ~9% (Figure B.7a-c and Figure B.S4a and c). These results recapitulate our previous observations with NMDAR antagonists (Workman et al., 2013). In addition, these findings in WT mouse neurons are consistent with what we observed in rat cultured hippocampal neurons treated with ethanol or the clinically relevant rapid antidepressant Ro 25-6981 (Figure B.S5). Unexpectedly, $GABA_BR$ activation in saline-treated Fmr1 KO neurons failed to reduce the calcium signal. Moreover, in ethanol-treated Fmr1 KO neurons, $GABA_BR$ activation failed to increase dendritic calcium signal (Figure B.7 b-c and Figure B.S4b and c). These findings suggest that the loss of FMRP in Fmr1 KO dendrites decouples $GABA_BRs$ from potassium channels. These results also suggest that the dynamic, ethanol-induced plasticity in $GABA_BR$ signaling, which is observed with rapid antidepressants, requires FMRP (Workman et al., 2013).

To further substantiate that FMRP regulates ethanol-dependent $GABA_BR$ plasticity, we overexpressed FMRP in rat hippocampal neurons. Overexpression of FMRP

Immunofluorescent images and quantification summary of $GABA_BR1$ surface expression in wild type (WT) and *Fmr1* KO primary hippocampal cultures normalized to MAP2 as volume control. **a**. **and b**. Representative images of immunostaining. **c**. Increased expression of surface $GABA_BR1$ in WT dendrites after ethanol (ETOH: 30 mM, 2 h) compared with vehicle (Veh: H_2O, 2 h) or ethanol–cycloheximide (30 mM ETOH+50 μM CHX, 2 h) treatment. No significant change in surface $GABA_BR1$ expression in *Fmr1* KO cultures treated with ETOH or ETOH+CHX was observed. WT surface $GABA_BR1$: Veh=1.00±0.08, *n*=28 dendrites; ETOH=1.76±0.17, *n*=37 dendrites; ETOH+CHX=1.22±0.1, *n*=36 dendrites. *Fmr1* KO surface $GABA_BR1$: Veh=1.36±0.13, *n*=39 dendrites; ETOH=1.44±0.15, *n*=42 dendrites; ETOH+CHX=0.91±0.09, *n*=29 dendrites. **d.**–**i.** BONCAT combined with PLA, a method to detect newly synthesized proteins. **d.**–**g.** Representative images for $GABA_BR1$ and $GABA_BR2$ expression. Pixels were equally dilated by 1 using ImageJ software for enhanced visualization as described by Cajigas *et al.* (Cajigas et al., 2012). In WT and *Fmr1* KO primary hippocampal cultures **h**. $GABA_BR1$ synthesis in dendrites was not altered by ethanol (30 mM, 2 h) compared with vehicle (H_2O, 2 h) treatment normalized to MAP2. WT $GABA_BR1$: Veh=1.00±0.09, *n*=47 dendrites; ETOH=0.82±0.08, *n*=39 dendrites. *Fmr1* KO $GABA_BR1$: Veh=0.64±0.09, *n*=38 dendrites; ETOH=0.48±0.05, *n*=36 dendrites. **i**. In contrast, ethanol induced a significant increase in new $GABA_BR2$ synthesis in WT hippocampal dendrites but not in *Fmr1* KO dendrites. WT $GABA_BR2$: Veh=1.00±0.06, *n*=21 dendrites; ETOH=1.39±0.11, *n*=25 dendrites. *Fmr1* KO $GABA_BR2$: Veh=1.48±0.10, *n*=32 dendrites; ETOH=1.23±0.07, *n*=41 dendrites. Significance determined by two-way analysis of variance with Tukey's post test. Values represent mean±s.e.m. Scale bars, 5 μm.

did not alter the $GABA_BR$ activation in saline-treated neurons because baclofen reduced the dendritic calcium signal. However, in ethanol-treated neurons, overexpressing FMRP blocked the ethanol-i nduced $GABA_BR$ plasticity (Figure B.7d and Figure B.S4d-e). These results provide additional evidence that the dynamic reduction of FMRP with ethanol exposure is important for the expression of $GABA_BR$ plasticity.

Antidepressant effect of ethanol on behavior requires FMRP

Since FMRP is important for ethanol-induced $GABA_BR$ plasticity, we examined antidepressant and anxiolytic-like effects of ethanol on behavior in Fmr1 KO mice. Interestingly, ethanol administration did not affect the behaviors of Fmr1 KO mice in the splash and open field tests compared to saline-treated mice (Figure B.S6a-e). Surprisingly, the basal state of immobility in the FST in Fmr1 KO mice is equivalent to ethanol-injected WT mice (Figure B.8). To explore this paradox, we examined the requirement of $GABA_BR$ activation in ethanol-induced decreases in immobility by using CGP-35348 to inhibit postsynaptic $GABA_BRs$. We previously showed that $GABA_BR$ antagonism blocked the antidepressant-like behavior produced by NMDAR antagonist in the FST (Workman et al., 2013). $GABA_BR$ inhibition alone did not affect the immobility of saline-injected WT mice in the FST, similar to what we observed previously (Figure B.8) (Workman et al., 2013). CGP-35348, however, abolished the ethanol-induced antidepressant behavior, demonstrating a requirement for $GABA_BR$ activation in ethanol-triggered reduction of immobility. Neither ethanol, CGP-35348, or ethanol + CGP-35348 treatment in Fmr1 KO mice produced immobility scores that were significantly different from saline-treated Fmr1 KO mice. These findings collectively demonstrate that $GABA_BRs$ and FMRP are necessary to elicit the ethanol-mediated antidepressant response.

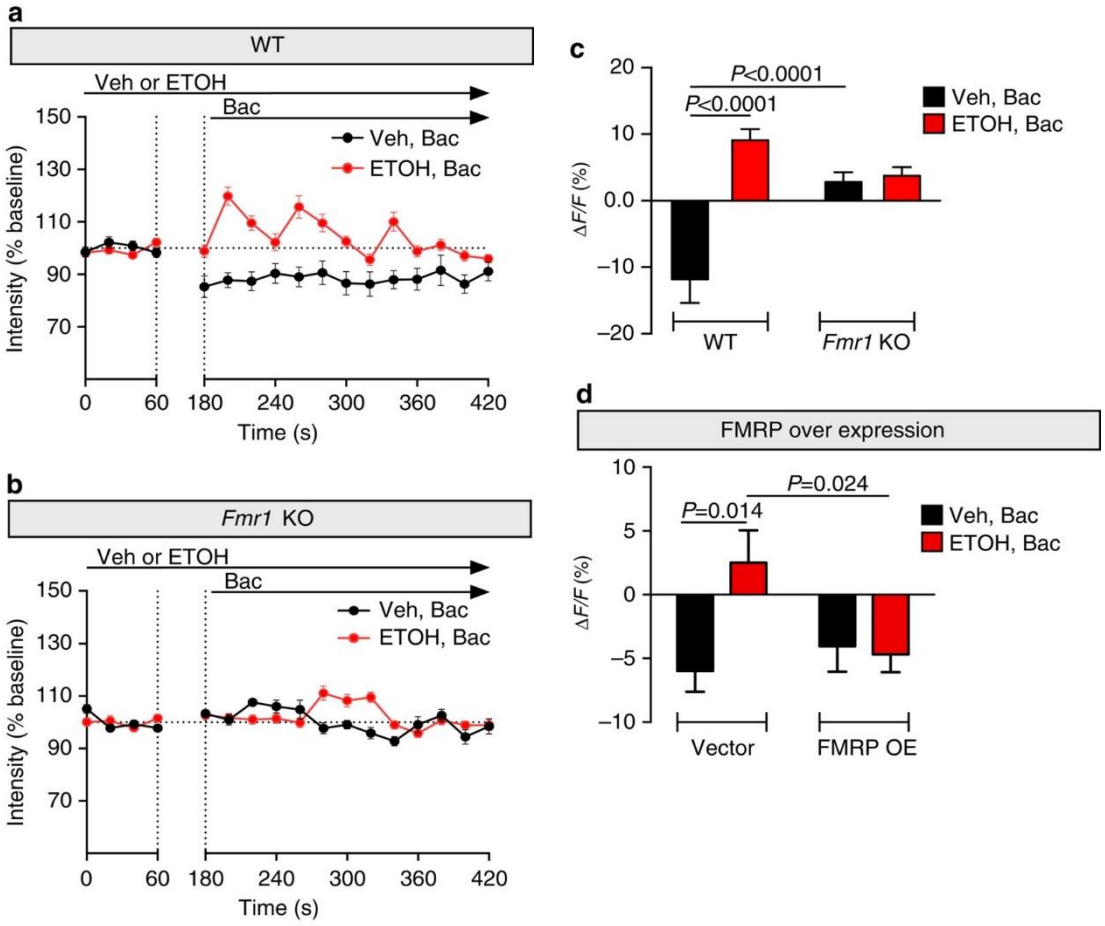


Figure B.7: GABA$_B$R plasticity and signaling is absent in *Fmr1* KO mice

a. and c. Mouse hippocampal cultured neurons were pre-treated for 2 h with either vehicle (Veh: H_2O) or ethanol (ETOH: 30 mM). Line graphs represent the average fluorescent calcium signal in dendrites over time from **a**. wild type (WT) and **b**. *Fmr1* KO mice. Baseline was established for 1 min before the addition of GABA$_B$R agonist baclofen (Bac: 50 μM) in vehicle- or ethanol-exposed neurons. Baclofen was allowed to equilibrate as indicated by the break between dotted lines. **c**. Summary graph shows significant increase in dendritic calcium signal ($\Delta F/F$) with the addition of baclofen in WT neurons pre-treated with ethanol, which was not observed in *Fmr1* KO neurons. WT: Veh+Bac= −11.82±3.55, n=8; ETOH+Bac=9.10±1.65, n=14. *Fmr1* KO: Veh+Bac=2.81±1.48, n=12; ETOH+Bac=3.74±1.30, n=12. **d**. Dendritic calcium imaging was performed as before in hippocampal cultured neurons infected with either vector (rAAV:mSYN-tdTomato) or FMRP overexpression (rAAV:mSYN-FMRP and rAAV:mSYN-tdTomato). Ethanol-induced increase in dendritic calcium is prevented by FMRP overexpression. Vector: Veh+Bac: −6±1.6, n=17 dendrites; ETOH+Bac: 2.5±2.5, n=17 dendrites. FMRP overexpression: Veh+Bac: −4±2, n=11 dendrites; ETOH+Bac: −4.7±1.4, n=27 dendrites. Significance determined by two-way analysis of variance, followed by Tukey's multiple comparison. Values represent mean±s.e.m.

DISCUSSION

Emerging behavioral and molecular evidence demonstrate that NMDAR antagonists act as rapid antidepressants (Li et al., 2010; Autry et al., 2011; Workman et al., 2015; Jaso et al., 2017). Because it has long been speculated that individuals with major depressive disorders self-medicate with alcohol, we examined whether ethanol, which blocks NMDARs (Lovinger et al., 1989), acts through the same synaptic pathways as NMDAR antagonists. Until this study, the molecular mechanisms shared by alcohol and antidepressants were unexplored. Here, we provide molecular and behavioral evidence that acute alcohol exposure elicits antidepressant-like behaviors that persist up to 24 hours after administration (Figure B.1), supporting the hypothesis that ethanol initiates lasting antidepressant activity. We have previously demonstrated that NMDAR inhibition by rapid antidepressants induces two key molecular changes that are responsible for the rapid antidepressant response, namely (1) an increase in $GABA_BR$ protein synthesis and (2) a shift in $GABA_BR$ function that increases dendritic calcium signaling (Workman et al., 2013; Workman et al., 2015). Our current work shows that these same signature changes are produced by acute ethanol exposure (Figure B.5-B.7).

Surface expression of functional $GABA_BRs$ requires the dimerization of $GABA_BR1$ and R2 subunits. Without $GABA_BR2$, $GABA_BR1$ is retained in the endoplasmic reticulum (Margeta-Mitrovic et al., 2000). Our current studies show that the release of $GABA_BR2$ mRNA translational repression by FMRP is necessary for the ethanol-induced increase in surface $GABA_BRs$ with NMDAR blockade (Figure B.4,B.6,B.8, B.S3). Reduction of FMRP, as seen in animal models of FXS, is associated with elevated protein synthesis of target mRNAs (Bassell and Warren, 2008). While we have demonstrated that FMRP associates with $GABA_BR1$ and R2 mRNAs, the loss of

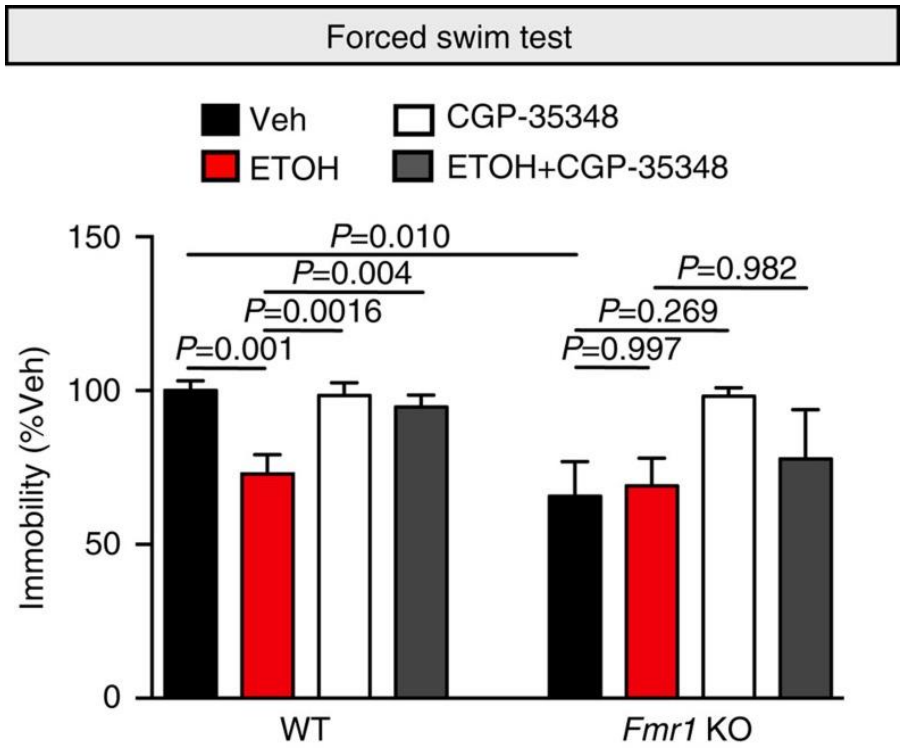


Figure B.8: Ethanol's antidepressant effect requires $GABA_BR$ activation

Wild type (WT) C57BL/6 and *Fmr1* KO male mice were subjected to the forced swim test (FST) 24 h post injection of vehicle (Veh: saline), ethanol (ETOH: 2.5 g kg^{-1}), CGP-35348, a GABABR antagonist (CGP: 100 mg kg^{-1}) or ethanol+CGP-5348. Ethanol-induced decrease in immobility was absent in *Fmr1* KO mice. WT: Veh=100±3.19 s, *n*=9 mice; ETOH=72.97±6.23 s, *n*=7 mice; CGP-35348=98.38±4.2 s, *n*=10; ETOH+CGP-35348=94.73±3.77 s, *n*=7 mice. *Fmr1* KO: Veh=58.75±10.33 s, *n*=9; ETOH=69.02±8.99 s, *n*=3; CGP-35348=88.00±9.56 s, *n*=3; ETOH+CGP-35348=77.78±16.04 s, *n*=3. Significance determined by two-way analysis of variance Tukey's multiple comparison test. Values represent mean±s.e.m.

FMRP has a profound effect on $GABA_BR2$ protein expression in dendrites. Constitutive loss of FMRP, as observed in many of its targets, abrogates stimulus-dependent mRNA transport and translation of target mRNAs (Muddashetty et al., 2007; Dictenberg et al., 2008; Soden and Chen, 2010; Niere et al., 2012). We show that the ethanol-induced increase in $GABA_BR2$ protein is also absent in Fmr1 KO mice.

The question of why $GABA_BR2$ is uniquely affected by acute ethanol and its dependence on reduced FMRP levels is intriguing. Both $GABA_BR1$ and $GABA_BR2$ mRNA were detected in the FMRP RNA-IP (Figure B.3). Interestingly, in vivo FMRP influences $GABA_BR1$ expression to a lesser extent compared to $GABA_BR2$, suggesting that FMRP may act in concert with other repressors such as microRNAs to tightly regulate $GABA_BR1$ expression (Muddashetty et al., 2011). Moreover, our immunostaining and protein synthesis assays suggest that the increased protein levels of $GABA_BR1$ are not due to new protein synthesis in the dendrites, even in Fmr1 KO mice (Figure B.5 and B.6). These results indicate that the overall increase in synaptic $GABA_BR1$ expression in Fmr1 KO mice may be due to an increase in presynaptic $GABA_BR1$ expression, either through protein synthesis or increased protein stability. Notably, FMRP has been localized to axons and presynaptic terminals (Antar et al., 2006; Christie et al., 2009). Further exploration into presynaptic and postsynaptic $GABA_BR$ expression and function in AUD and FXS is warranted.

The "GABA hypoinhibition theory" posits that loss of inhibition is a leading cause in many of the neurological symptoms observed in FXS (Gatto et al., 2014). While studies showing reduced inhibition in models of FXS have focused on decreased expression of $GABA_BR$ subunit mRNA and protein (D'Hulst et al., 2006), $GABA_BR$ protein expression and dendritic signaling has not been explored. Interestingly, the $GABA_BR$ agonist baclofen has shown promise in treating FXS. (R)-baclofen administration in vitro corrects the

elevated basal protein synthesis normally seen in Fmr1 KO mice, and rescues synaptic abnormalities such as increased spine density (Henderson et al., 2012). Additionally, baclofen administration reduced symptoms related to FXS in Fmr1 KO mice (Henderson et al., 2012; Qin et al., 2015). Recent studies have shown that the excitatory drive to fast spiking inhibitory neurons is reduced in the cortex of Fmr1 KO mice (Gibson et al., 2008). Our data expands upon these findings by suggesting that postsynaptic $GABA_BR$ coupling to inwardly rectifying G protein-coupled potassium channels (GIRK) is absent in Fmr1 KO neurons (Figure B.7). Collectively, these results may imply that the therapeutic effects of baclofen in the Fmr1 KO mouse may be due to activation of presynaptic $GABA_B$ receptors that may in turn reduce glutamate release and reduce hyperactive mGluR signaling in the hippocampus (Osterweil et al., 2012).

Changes in gene expression and protein synthesis are essential for normal neuroplasticity, but these crucial processes are dysregulated by drug addiction (Ron and Messing, 2013; Kenny, 2014; Folsom et al., 2015; Fuchsova et al., 2015). Several lines of evidence support parallel changes in $GABA_BR$ mRNA translation/signaling as a result of NMDAR blockade that may be critical for alcohol actions. First, NMDAR antagonists mimic some effects of ethanol in humans (Holmes et al., 2013), suggesting common biochemical/electrophysiological signaling pathway(s). Second, changes in $GABA_BR2$ brain gene expression correlates with lifetime alcohol consumption, supporting a role for altered $GABA_BR$ signaling in AUD (Farris et al., 2015). Third, although controversial, the $GABA_BR$ agonist baclofen may decrease alcohol consumption in some alcoholics (Krishnan-Sarin et al., 2008). In summary, our data defines a common molecular paradigm for alcohol and rapid antidepressants and identifies a mechanism for the initial antidepressant effects of alcohol. A shift in $GABA_BR$ signaling is observed with both rapid

antidepressants and acute ethanol treatment, which may provide insight into the molecular basis for the high comorbidity between major depressive disorder and AUD.

Appendix C: Code for Chapters 2 and 3[13]

Code for Data Processing

This section includes all the code used to process the normalized spectral counts. The raw data (normalized spectral counts) can be found in Table S1A (available online). Copying the data from the cells B12:T756 will give you what you need to replicate this script exactly and produce the tidy data used to generate the figures seen in the paper.

First, load the plyr package:

```
library(plyr)
```

Load the data from your working directory.

```
raw <- read.csv("RawData.csv", header = TRUE, stringsAsFactors = FALSE)
```

Cosmetic processing. Name the rows, clean up a bit, swap 0's to NA.

```
names(raw) <- c("Protein", "LD1", "LD2", "LD3",
                           "LR1", "LR2", "LR3",
                           "PD1", "PD2", "PD3",
                           "PR1", "PR2", "PR3",
                           "SD1", "SD2", "SD3",
                           "SR1", "SR2", "SR3")

rownames(raw) <- raw[, "Protein"]
raw <- raw[, -1]
raw[raw == 0] <- NA   # Changing NAs to zero to help with processing do
wnstream
```

[13] This code and additional files deposited at https://github.com/snamjoshi/MCP2015_Niere_Namjoshi_etal

Split the table into separate data frames for rapa/DMSO

```
lysDMSO <- raw[, 1:3]
lysRapa <- raw[, 4:6]
psdDMSO <- raw[, 7:9]
psdRapa <- raw[, 10:12]
solDMSO <- raw[, 13:15]
solRapa <- raw[, 16:18]
```

Next, calculate the number of proteins that returned 0 spectral counts for each replicate. Remove the proteins from the data frame if they have 0 values for 2 or 3 of 3 replicates. We consider these measurements to be unreliable.

```
lysDMSO <- lysDMSO[which(apply(lysDMSO, 1, function(x) sum(is.na(x))) <
= 1), ]
lysRapa <- lysRapa[which(apply(lysRapa, 1, function(x) sum(is.na(x))) <
= 1), ]
psdDMSO <- psdDMSO[which(apply(psdDMSO, 1, function(x) sum(is.na(x))) <
= 1), ]
psdRapa <- psdRapa[which(apply(psdRapa, 1, function(x) sum(is.na(x))) <
= 1), ]
solDMSO <- solDMSO[which(apply(solDMSO, 1, function(x) sum(is.na(x))) <
= 1), ]
solRapa <- solRapa[which(apply(solRapa, 1, function(x) sum(is.na(x))) <
= 1), ]
```

We now have a filtered list of proteins by fraction and treatment. Looking at all the proteins, we see that these numbers correspond to what we reported in the publication.

```
##   Fraction Proteins
## 1  lysDMSO      506
## 2  lysRapa      504
## 3  psdDMSO      459
## 4  psdRapa      449
## 5  solDMSO      548
## 6  solRapa      552
```

Now find the row means for all fractions/treatments (to average the replicates)

```
meanLysDMSO <- rowMeans(lysDMSO, na.rm = TRUE)
meanLysRapa <- rowMeans(lysRapa, na.rm = TRUE)
meanPsdDMSO <- rowMeans(psdDMSO, na.rm = TRUE)
meanPsdRapa <- rowMeans(psdRapa, na.rm = TRUE)
meanSolDMSO <- rowMeans(solDMSO, na.rm = TRUE)
meanSolRapa <- rowMeans(solRapa, na.rm = TRUE)
```

Next we want to combine rapa and DMSO into one data frame where NA indicates that proteins was not found in that fraction/treatment. To do this we need to create a function that takes the DMSO data frame and the rapa data frame and combines them together. The `mergeVec` function below will accomplish this.

```
mergeVec <- function(x, y) {
    require(plyr)

    input <- list(as.data.frame(t(x)), as.data.frame(t(y)))
    input <- t(do.call(rbind.fill, input))

    return(input)
}
```

Now apply the `mergeVec function` to the data. We will add column names and change NA values back to 0.

```
lysate <- mergeVec(meanLysDMSO, meanLysRapa)
psd <- mergeVec(meanPsdDMSO, meanPsdRapa)
soluble <- mergeVec(meanSolDMSO, meanSolRapa)

colnames(lysate) <- c("DMSO", "Rapa")
colnames(psd) <- c("DMSO", "Rapa")
colnames(soluble) <- c("DMSO", "Rapa")

lysate[is.na(lysate)] <- 0
psd[is.na(psd)] <- 0
soluble[is.na(soluble)] <- 0
```

Next we want to compute the fold changes for each fraction. This will require a function that divides the Rapa value by the DMSO value and then transforms the data using $\log_2$. If the output returns Inf (divide by zero) or -Inf ($\log_2(0)$), then the output will be assigned the characters “HIGH” or “LOW” respectively.

The `foldChange` function below will accomplish this.

```
foldChange <- function(fraction) {
    output <- as.data.frame(fraction[, 2] / fraction[, 1])
    output <- log2(output)

    output[output == Inf, ] <- "HIGH"
    output[output == -Inf, ] <- "LOW"

    return(output)
}
```

Apply the `foldChange` function to the data.

```
lysateFC <- foldChange(lysate)
psdFC <- foldChange(psd)
solubleFC <- foldChange(soluble)
```

Now merge all the lists together for each fraction using the `mergeVec` function. We do this in two steps: first for lysate and PSD and finally adding in the soluble column.

```
fullList <- as.data.frame(mergeVec(lysateFC, psdFC))
fullList <- as.data.frame(mergeVec(fullList, solubleFC))
```

Last bit of ordering and cosmetic changes to get the data frame into a neat format.

```
fullList <- fullList[ order(row.names(fullList)), ]
fullList <- cbind(row.names(fullList), fullList)
row.names(fullList) <- 1:length(fullList[, 1])
colnames(fullList) <- c("Gene", "Lysate", "PSD", "Soluble")
```

This produces the full list of proteins, our processed data. The final data frame here (or in some cases the precursor data frames above) will be used downstream for other analysis found in the paper.

CODE FOR FIGURE 2.2

To use this script you need the data generated from DataProcessing.R. In that document this data was assigned to the variable "fullList".

First, load the libraries.

```
library(ggplot2)
library(reshape2)
library(moments)
```

Figure 2.2D

Import data.

```
fullList <- read.csv("fullList.csv", header = TRUE)
```

Cosmetic changes and get data into long form. "HIGH" and "LOW" are changed to NA.

```
fullList <- fullList[, -1]
fullList[fullList == "HIGH"] <- NA
fullList[fullList == "LOW"] <- NA
fullList <- apply(fullList, 2, as.numeric)
fullList <- melt(fullList)
fullList <- fullList[, -1]
colnames(fullList) <- c("Fraction", "Expression")
```

Subset the PSD fraction and generate the graph.

```
histLongPSD <- subset(fullList, Fraction == "PSD")

binsize <- diff(range(na.omit(histLongPSD$Expression)))/50
```

```
a <- ggplot(histLongPSD, aes(x = Expression, fill = Fraction)) +
    geom_histogram(position = "identity", binwidth = binsize) +
    geom_vline(xintercept = 0, linetype = "dashed", color = "black", si
ze = 0.3) +
    geom_vline(xintercept = 0.1375, linetype = "solid", color = "red",
size = 0.3) +
    geom_vline(xintercept = -0.1375, linetype = "solid", color = "red",
size = 0.3) +
    scale_fill_manual(values = c("black")) +
    scale_x_continuous(limits = c(-1.75, 1.75), breaks = seq(-2, 2, by
= 0.5)) +
    xlab("Log2 Fold Change") +
    ylab("Number of Proteins") +
    theme_bw() +
    theme(panel.grid.major = element_blank(), panel.grid.minor = elemen
t_blank(),
                panel.background = element_blank(), axis.line = element
_line(color = "black"))

a +
    theme(text = element_text(size = 15)) +
    theme(legend.position = "none") +
    annotate("rect", xmin = -0.136, xmax = 0.1375, ymin = 0, ymax = Inf
, fill = "grey99", alpha = 0.9)
```

Statistics in Figure 2D Legend

Import the data sets needed.

```
fullList <- read.csv("fullList.csv", header = TRUE)
fullListPSD <- fullList[ ,c(1,3)]
fullListPSD$PSD <- as.numeric(as.character(fullListPSD$PSD))
```

```
## Warning: NAs introduced by coercion
```

Number of proteins calculations (excludes out-of-range proteins):

```
sum(fullListPSD$PSD > 0, na.rm = TRUE)
```

```
## [1] 210
```

```
sum(fullListPSD$PSD < 0, na.rm = TRUE)

## [1] 191

sum(complete.cases(fullListPSD$PSD))

## [1] 401
```

This yields 210 PSD proteins with fold changes greater than 0, 191 PSD proteins with fold changes less than 0 and 401 total PSD proteins as reported.

Descriptive statistics reported in figure legend:

```
mean(fullListPSD$PSD, na.rm = TRUE) # Mean

## [1] -0.006047828

sd(fullListPSD$PSD, na.rm = TRUE) # Standard deviation

## [1] 0.4950849

var(fullListPSD$PSD, na.rm = TRUE) # Variance (CI calculated in GraphPa
d)

## [1] 0.2451091

skewness(fullListPSD$PSD, na.rm = TRUE) # Skewness

## [1] -0.39867

kurtosis(fullListPSD$PSD, na.rm = TRUE) # Kurtosis

## [1] 4.136472
```

Bartlett's test for comparison of variance in fold-change across fractions.

```
# Import Data and process
fullList <- read.csv("fullList.csv", header = TRUE)
fullList <- fullList[, -1]
fullList[fullList == "HIGH"] <- NA
fullList[fullList == "LOW"] <- NA
fullList <- apply(fullList, 2, as.numeric)
fullList <- melt(fullList)
```

```
fullList <- fullList[, -1]
colnames(fullList) <- c("Fraction", "Expression")
fullList <- fullList[complete.cases(fullList), ]

# Compute Bartlett's Test
bartlett.test(Expression~Fraction, fullList) # p < 0.0002 | Variances a
re different

##
##  Bartlett test of homogeneity of variances
##
## data:  Expression by Fraction
## Bartlett's K-squared = 17.1834, df = 2, p-value = 0.0001856
```

Figure 2.2E

Import the data.

```
fullList <- read.csv("fullList.csv", header = TRUE, stringsAsFactors =
FALSE)

nrow(subset(fullList, fullList$PSD == "HIGH", select = c("Gene", "PSD")
))

## [1] 48

nrow(subset(fullList, fullList$PSD == "LOW", select = c("Gene", "PSD"))
)

## [1] 58
```

However, note that for the out-of-range genes we used an even more stringent classification. If either treatment had 2 zero values, we would not classify it as out-of-range because we could not reliably compute the fold-change. These were manually removed after the script was written and thus shortens the list. The graph shown in figure 2E was generated in GraphPad.

CODE FOR FIGURE 2.3

To use this code you will need the proteins in Table S7A (online) (cells A12:C615) and "fullList.csv" which was generated in DataProcessing.R. You will also need the list of mitchondrial genes in "mito.csv".

Load libraries and data lists.

```
library(gplots)

# Disease lists
disease <- read.csv("diseaseCandidates.csv", header = TRUE, stringsAsFa
ctors = FALSE, na.strings = "")

diseaseList <- list(Epilepsy = disease$Epilepsy,
                                        Alzheimers = disease$Alzheimers
,
                                        ASD = disease$ASD)

diseaseList <- lapply(diseaseList, na.omit)
diseaseList <- lapply(diseaseList, as.data.frame) # Removes na.action a
ttributes

# Mass spec list
fullList <- read.csv("fullList.csv", header = TRUE)
fullList$Gene <- toupper(fullList$Gene)  # Supp table is in uppercase,
easier comparison

# Mitochondrial gene list
mito <- read.csv("mito.csv", header = FALSE, stringsAsFactors = FALSE)
mito <- toupper(mito[ ,1])
```

Figure 2.3A

Process data and generate heatmap for epilepsy genes.

```
epilepsy <- diseaseList$'Epilepsy'
epilepsy <- fullList[fullList$Gene %in% epilepsy[ ,1], ] # Filter to is
olate genes in our dataset
epilepsy <- epilepsy[!epilepsy$Gene %in% mito, ]
epilepsy <- sapply(epilepsy[ ,2:4], function(x) as.numeric(as.character
(x)))
```

```
epilepsyMatrix <- data.matrix (epilepsy)

# Labels and out-of-range gene textured boxes added manually in Inkscap
e later
heatmap.2(
    as.matrix(epilepsyMatrix[ ,c(2,3)]),
    dendrogram = "none",
    Rowv = NULL,
    Colv = NULL,
    col = "heat.colors",
    trace = "none",
    na.color = "grey",
    density.info = "none",
    #key = TRUE,
    #keysize = 0.5,
    labCol = FALSE
)
```

Figure 2.3B

Process data and generate heatmap for Alzheimer's genes

```
alz <- diseaseList$'Alzheimers'
alz <- fullList[fullList$Gene %in% alz[ ,1], ] # Filter to isolate gene
s in our dataset
alz <- alz[!alz$Gene %in% mito, ]
alz <- sapply(alz[ ,2:4], function(x) as.numeric(as.character(x)))
alzMatrix <- data.matrix(alz)

# Labels and out-of-range gene textured boxes added manually in Inkscap
e later
heatmap.2(
    as.matrix(alzMatrix[ ,c(2,3)]),
    dendrogram = "none",
    Rowv = NULL,
    Colv = NULL,
    col = "heat.colors",
    trace = "none",
    na.color = "grey",
    density.info = "none",
    #key = TRUE,
    #keysize = 0.5,
    labCol = FALSE
)
```

Figure 2.3C

Process data and generate heatmap for ASD genes

```
asd <- diseaseList$'ASD'
asd <- fullList[fullList$Gene %in% asd[ ,1], ] # Filter to isolate gene
s in our dataset
asd <- asd[!asd$Gene %in% mito, ]
asd <- sapply(asd[ ,2:4], function(x) as.numeric(as.character(x)))
asdMatrix <- data.matrix(asd)

# Labels and out-of-range gene textured boxes added manually in Inkscap
e later
heatmap.2(
    as.matrix(asdMatrix[ ,c(2,3)]),
    dendrogram = "none",
    Rowv = NULL,
    Colv = NULL,
    col = "heat.colors",
    trace = "none",
    na.color = "grey",
    density.info = "none",
    #key = TRUE,
    #keysize = 0.5,
    labCol = FALSE
)
```

CODE FOR FIGURE 3.6

This code utilizes material from various disease databases (collected in diseaseCandidates.csv), the fullList.csv file generated in the data processing step, and a disease list which includes the proteins from out mass spec data that overlap with the candidate proteins.

Load libraries and process data

```
library(VennDiagram)
```

Import data sets and process:

```
# Disease lists
disease <- read.csv("diseaseCandidates.csv", header = TRUE, stringsAsFa
ctors = FALSE, na.strings = "")

diseaseList <- list(Epilepsy = disease$Epilepsy,
                                        Alzheimers = disease$Alzheimers
,
                                        ASD = disease$ASD)

diseaseList <- lapply(diseaseList, na.omit)

# Mass spec list
fullList <- read.csv("fullList.csv", header = TRUE)
fullList$Gene <- toupper(fullList$Gene)  # Supp table is in uppercase,
easier comparison

# Epilepsy List
epilepsy <- diseaseList$'Epilepsy'
epilepsy <- fullList[fullList$Gene %in% epilepsy, ] # Filter to isolate
genes in our dataset

# Alzheimer's List
alz <- diseaseList$'Alzheimers'
alz <- fullList[fullList$Gene %in% alz, ] # Filter to isolate genes in
our dataset

# ASD List
asd <- diseaseList$'ASD'
asd <- fullList[fullList$Gene %in% asd, ] # Filter to isolate genes in
our dataset
```

The VennSet function returns a list of each element of a 3-way Venn diagram. I am not the original author of this function; I believe I found this function in a StackExchange post. Please notify me if you come across the original source so I can give credit to the author.

```
VennSet <- function(A, B, C)
{
    unionAB <- union(A, B)
    unionAC <- union(A, C)
    unionBC <- union(B, C)
    uniqueA <- setdiff(A, unionBC)
    uniqueB <- setdiff(B, unionAC)
    uniqueC <- setdiff(C, unionAB)
    intersAB <- setdiff(intersect(A, B), C)
    intersAC <- setdiff(intersect(A, C), B)
    intersBC <- setdiff(intersect(B, C), A)
    intersABC <- intersect(intersect(A, B), intersect(B, C))
    items <- list(A=uniqueA, B=uniqueB, C=uniqueC, AB=intersAB , AC=int
ersAC , BC=intersBC , ABC=intersABC)
    return(items)
}
```

Figure 3.6A (*left*)

Find the different intersections and determine the number of elements in each. This is what will be fed into the Venn diagram.

```
diseaseVenn <- VennSet(diseaseList$Epilepsy, diseaseList$Alzheimers, di
seaseList$ASD)
lapply(diseaseVenn, length)
```

```
## $A
## [1] 319
##
## $B
## [1] 184
##
## $C
## [1] 532
##
## $AB
## [1] 11
##
## $AC
## [1] 57
##
## $BC
## [1] 12
##
## $ABC
## [1] 2
```

Draw the Venn diagram. In the paper the numbers and shapes for the Venn diagram were entered manually. I have reproduced them here with R code.

```
grid.newpage()
draw.triple.venn(area1 = 389, area2 = 209, area3 = 603, n12 = 13, n23 = 14, n13 = 59,
                                 n123 = 2, category = c("Epilepsy", "AD", "ASD"), lty = "blank",
                                 fill = c("brown1", "darkorchid2", "lightslateblue"))
```

Figure 3.6B (*right*)

First we need to import the list of disease candidates found in our mass spec data

```
OurDiseaseList <- read.csv("Diseases.csv", header = TRUE)
OurDiseaseList.list <- list(A = OurDiseaseList[,1], B = OurDiseaseList[,2], C = OurDiseaseList[,3])
```

Find the intersections and determine the number of elements in each.

```
disease.Venn <- VennSet(OurDiseaseList.list$A, OurDiseaseList.list$B, OurDiseaseList.list$C)
lapply(disease.Venn, length)
```

```
## $A
## [1] 40
##
## $B
## [1] 40
##
## $C
## [1] 13
##
## $AB
## [1] 3
##
## $AC
## [1] 5
```

```
##
## $BC
## [1] 1
##
## $ABC
## [1] 1
```

Draw the Venn diagram.

```
grid.newpage()
draw.triple.venn(area1 = 48, area2 = 44, area3 = 19, n12 = 3, n23 = 1,
n13 = 5,
                                 n123 = 0, category = c("Epilepsy", "AD
", "ASD"), lty = "blank",
                                 fill = c("rosybrown2", "plum", "skyblu
e1"))
```

CODE FOR SUPPLEMENTAL FIGURES

This document includes the code needed to calculate the coefficient of determination between samples of fractions and the measurement of error between the fractions. A correlation matrix is also included though this was not reported in the publication.

Data processing and package loading

First, load the packages.

```
library(ggplot2)
library(reshape2)
```

Load the data and get into tidy format. RawData.csv is the data in Supplemental Table 1A (online) from cells B12:T756. The only difference is that the rows have been sorted alphabetically.

```
raw <- read.csv("RawData.csv", header = TRUE, stringsAsFactors = FALSE)

names(raw) <- c("Protein", "LD1", "LD2", "LD3",
                                "LR1", "LR2", "LR3",
                                "PD1", "PD2", "PD3",
                                "PR1", "PR2", "PR3",
                                "SD1", "SD2", "SD3",
                                "SR1", "SR2", "SR3")

rownames(raw) <- raw[, "Protein"]
raw <- raw[, -1]
```

The `prop()` function will be used to determine the variance across the different proteins (reliability of spectral counts). The function first calculates the standard error among the proteins by treatment. These standard errors are then normalized by the mean spectral count for the protein. This gives the mean standard error for all proteins by treatment.

```
prop <- function(cols){
    x <- raw[, cols]
    x <- x[rowSums(x == 0) <= 1, ]
    x[x == 0] <- NA

    x.var <- apply(x, 1, function(r) {
        sd(r, na.rm = TRUE) / sqrt(sum(!is.na(r)))
    })

    x.prop <- x.var / rowMeans(x, na.rm = TRUE)
    return(x.prop)
}
```

The `prop()` function is then applied to all fractions.

```
totals <- list(LD = prop(1:3),
                               LR = prop(4:6),
                               PD = prop(7:9),
                               PR = prop(10:12),
                               SD = prop(13:15),
                               SR = prop(16:18))
```

Bar Charts for figures 2.6 and 2.7

A standard error function for use below.

```
se <- function(x) sqrt(var(x)/length(x))
```

Arrange the data for graphing by calculating the mean and standard deviation and putting them into a data frame.

```
totals.mean <- sapply(totals, mean)
totals.se <- sapply(totals, se)
totals.bar <- data.frame(Mean = totals.mean, SE = totals.se)
```

For the publication, we put these values into GraphPad. The graph can be replicated in R with the following code:

```
ggplot(totals.bar, aes(x = rownames(totals.bar), y = Mean, fill = rowna
mes(totals.bar), color = rownames(totals.bar))) +
    geom_bar(stat = "identity", color = "black", fill = "grey") +
    geom_errorbar(aes(ymin = Mean + 0, ymax = Mean + SE), width = 0.2,
color = "black") +
    xlab("Fraction") +
    ylab("Normalized Mean S.E.M.") +
    coord_cartesian(ylim = c(0,0.25)) +
    theme_bw()
```

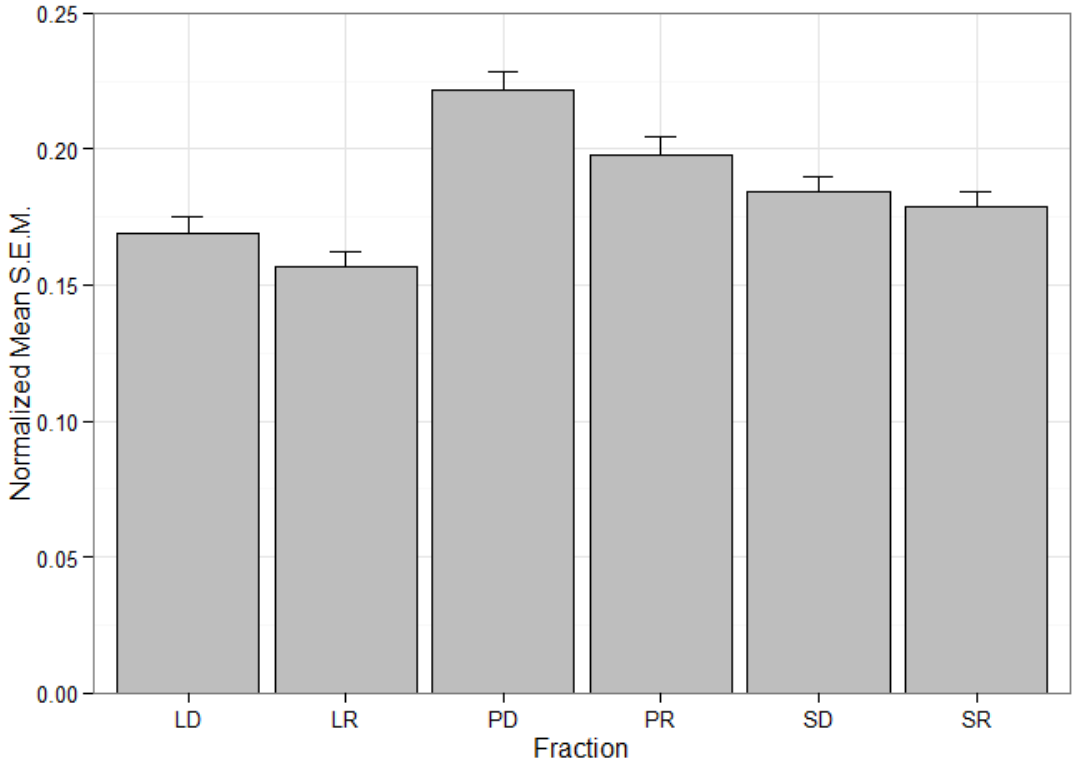


Figure C.1: Average S.E.M. of the normalized spectral counts across all genes within each fraction

The following is not included in the paper but may be useful. We can plot the probability density functions for each of the fractions and see how the distribution of standard errors compares.

```
totals.long <- melt(totals)

ggplot(totals.long, aes(x = value, color = L1)) +
    geom_density() +
    xlab("Distribution of S.E.M.") +
    ylab("Density") +
    theme_bw()
```

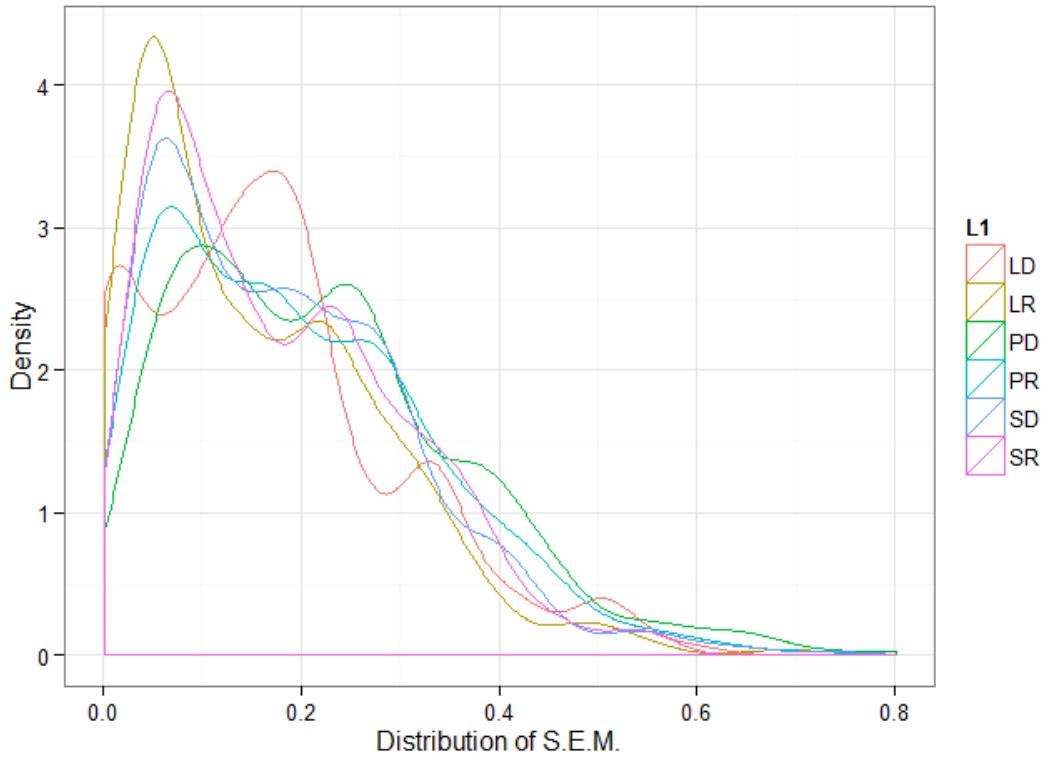


Figure C.2: Probability distribution of standard errors across all genes within each fraction

Correlations for figures 2.6 and 2.7

Here we use the coefficient of determination to see the correlation of the triplicate measurements among all normalized spectral counts by treatment and fraction. We will still use the raw data loaded above for the analysis (stored in the variable `raw`).

The correlations need to be done in sets of three. We can create the pattern sequence we need for each set with the code displayed below. We will store the pattern key-index in the variable `m`.

```
v1 <- c(1:18)
v2 <- v1+1L
v3 <- c(0,0,3L)
```

```
v2 <- v2-v3
m <- cbind(v1,v2)
```

We apply the `lm()` function to the raw data and extracted `r.squared` in sets of three. The key specifies the pattern in which the correlations are to be computed. We assign the correlations to the variable `fit`.

```
fit <- lapply(1:length(raw),function(x) summary(lm(raw[,m[x,1]]~raw[,m[
x,2]]))$r.squared)
```

Now we just have to make some cosmetic adjustments to make fit look tidy.

```
data.frame(Fraction = c(rep("Lysate", 6), rep("PSD", 6), rep("Soluble",
6)),
                    Treatment = rep(c("DMSO","RAPA"), 3, each = 3),
                    Replicate = rep(c("1 vs 2", "2 vs 3", "3 vs 1"), 6
),
                    R2 = round(unlist(fit), 3))
```

```
##    Fraction Treatment Replicate    R2
## 1    Lysate      DMSO    1 vs 2 0.922
## 2    Lysate      DMSO    2 vs 3 0.959
## 3    Lysate      DMSO    3 vs 1 0.900
## 4    Lysate      RAPA    1 vs 2 0.943
## 5    Lysate      RAPA    2 vs 3 0.949
## 6    Lysate      RAPA    3 vs 1 0.931
## 7       PSD      DMSO    1 vs 2 0.794
## 8       PSD      DMSO    2 vs 3 0.898
## 9       PSD      DMSO    3 vs 1 0.856
## 10      PSD      RAPA    1 vs 2 0.855
## 11      PSD      RAPA    2 vs 3 0.905
## 12      PSD      RAPA    3 vs 1 0.867
## 13  Soluble      DMSO    1 vs 2 0.832
## 14  Soluble      DMSO    2 vs 3 0.900
## 15  Soluble      DMSO    3 vs 1 0.907
## 16  Soluble      RAPA    1 vs 2 0.900
## 17  Soluble      RAPA    2 vs 3 0.941
## 18  Soluble      RAPA    3 vs 1 0.843
```

Additionally, we can also look at a correlation matrix. This does not appear in the publication but may also be useful.

Find correlations across the entire data.

```
corDat <- cor(raw)
corDatMelt <- melt(corDat)
```

To make the graph more readable we have scaled the colors in the correlation matrix across the range of R squared values (rather than scaled over 0 to 1). This makes the color shading much more readable visually. First we create a midpoint function. Then we find the midpoint of the correlations and set it to `mdpt`. This is then set as the midpoint of the color gradient in the `scale_fill_gradient2()` function call.

```
# Function to calculate midpoint for use in color scale
midpoint <- function(x) {
    ((max(x) - min(x))/2) + min(x)
}

# Set midpoint of data to variable mdpt
mdpt <- midpoint(corDatMelt$value)

# Plot correlation matrix
ggplot(corDatMelt, aes(x = Var1, y = Var2, fill = value)) +
    geom_tile() +
    scale_fill_gradient2(low = "red", mid = "orange", high = "white", m
idpoint = mdpt) +
    theme(axis.text.x = element_text(angle = 75, hjust = 1, vjust = 1))
+
    xlab("") +
    ylab("")
```

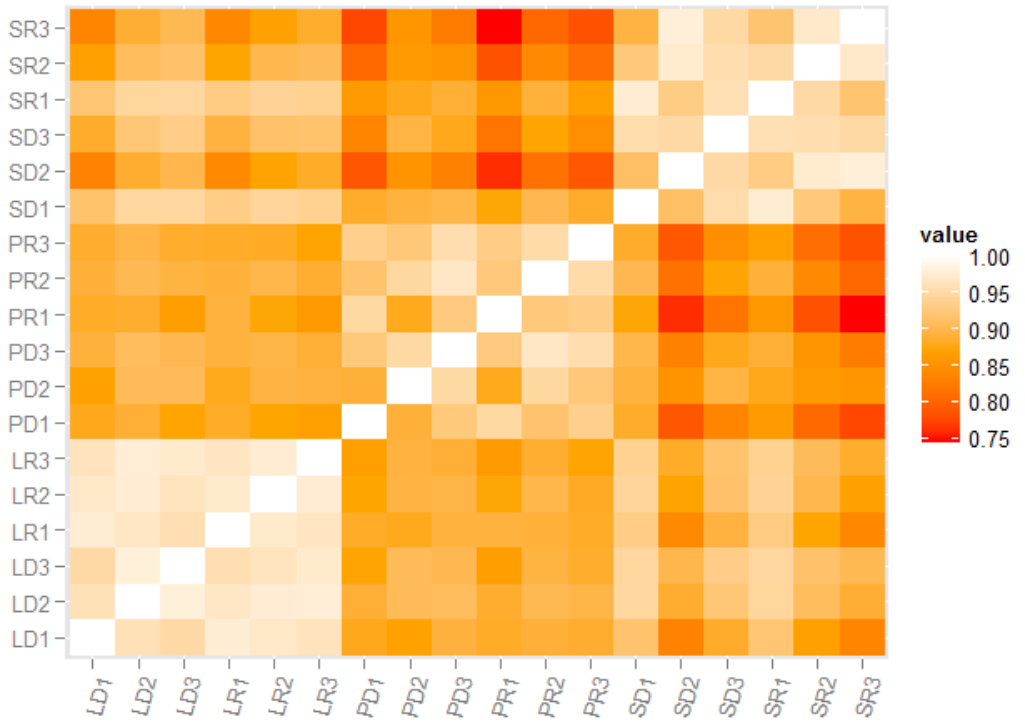


Figure C.3: Correlation matrix of all replicates in all fractions

CODE FOR SUPPLEMENTAL TABLES

This section includes all of the code and data files needed to replicate some of the results obtained from the online supplemental tables.

Table S2

First, let's create the `read.csv.list` function which reads a CSV file with uneven columns and converts it to a list instead of a dataframe. This is pretty common when uploading gene lists and this is an easy fix for it.

```
# Function reads a CSV file with uneven columns and coverts to a list i
nstead of a dataframe

read.csv.list <- function(file, header) {
    outputList <- apply(read.csv(file, header = header), 2, list)
```

```
    outputList <- lapply(outputList, function(x) lapply(x, '[', which(c
omplete.cases(x))))

    return(outputList)
}
```

Load the data. This data is a network stats file that has all of the network stats calculated by Cytoscape for the DMSO network, the rapa network, and a randomly generated network.

```
network <- read.csv.list("networkStats.csv", TRUE)
```

First we need to summarize the data by mean and standard deviation as it is included in the table.

```
meanLysate <- unlist(lapply(network, function(x) mean(as.numeric(unlist
(x)))))
sdLysate <- unlist(lapply(network, function(x) sd(as.numeric(unlist(x))
)))

summaryLysate <- data.frame(mean = round(meanLysate, 3), sd = round(sdL
ysate, 2))

summaryLysate
```

```
##                                    mean   sd
## DMSO_Closeness_Centrality         0.319 0.17
## Rapa_Closeness_Centrality         0.326 0.19
## Random_Closeness_Centrality       0.421 0.06
## DMSO_Clustering_Coefficient       0.735 0.16
## Rapa_Clustering_Coefficient       0.739 0.16
## Random_Clustering_Coefficient     0.127 0.02
## DMSO_Neighborhood_Connectivity   16.458 8.56
## Rapa_Neighborhood_Connectivity   15.881 7.32
## Random_Neighborhood_Connectivity 20.268 1.34
```

Next we perform all of the tests whose P-values are listed in the table. The DMSO network, rapa network, and random networks were generated in Cytoscape first and all of the network statistics data was exported.

```
# DMSO Lysate Analysis
dmsoAnalysis <- lapply(network[c(1,4,7)], function(x) wilcox.test(as.nu
meric(unlist(x)), alternative = "greater"))

# Rapa Lysate Analysis
rapaAnalysis <- lapply(network[c(2,5,8)], function(x) wilcox.test(as.nu
meric(unlist(x)), alternative = "greater"))

# Random Clustering Coefficent DMSO
randDmsoCC <- ks.test(as.numeric(unlist(network[[4]])), as.numeric(unli
st(network[[6]])), alternative = "two.sided")

# Random Clustering Coefficent Rapa
randRapaCC <- ks.test(as.numeric(unlist(network[[5]])), as.numeric(unli
st(network[[6]])), alternative = "two.sided")

# Random Neighborhood Connectivity DMSO
randDmsoNC <- ks.test(as.numeric(unlist(network[[7]])), as.numeric(unli
st(network[[9]])), alternative = "two.sided")

# Random Neighborhood Connectivity Rapa
randRapaNC <- ks.test(as.numeric(unlist(network[[8]])), as.numeric(unli
st(network[[9]])), alternative = "two.sided")
```

Next, we put it all into a table.

```
summary <- data.frame(P.value = c(unname(unlist(sapply(dmsoAnalysis, "[
", 3))), unname(unlist(sapply(rapaAnalysis, "[", 3))), randDmsoCC$p.val
ue, randRapaCC$p.value, randDmsoNC$p.value, randRapaNC$p.value),
                                                row.names = c("DMSO Closene
ss Centrality", "DMSO Clustering Coefficient", "DMSO Neighborhood Conne
ctivity", "Rapa Closeness Centrality", "Rapa Clustering Coefficient", "
Rapa Neighborhood Connectivity", "Random DMSO Clustering Coefficent", "
Random Rapa Clustering Coefficient", "Random DMSO Neighborhood Connecti
vity", "Random Rapa Neighborhood Connectivity"))

summary <- cbind(summary, LessThan0.001 = summary$P.value < 0.001)
```

```
summary
```

```
##                                          P.value LessThan0.001
## DMSO Closeness Centrality             1.768067e-38          TRUE
## DMSO Clustering Coefficient           1.862645e-09          TRUE
## DMSO Neighborhood Connectivity        9.313226e-10          TRUE
## Rapa Closeness Centrality             5.078745e-36          TRUE
## Rapa Clustering Coefficient           2.000189e-06          TRUE
## Rapa Neighborhood Connectivity        1.862645e-09          TRUE
## Random DMSO Clustering Coefficent     5.087042e-13          TRUE
## Random Rapa Clustering Coefficient    8.461010e-13          TRUE
## Random DMSO Neighborhood Connectivity 1.275006e-06          TRUE
## Random Rapa Neighborhood Connectivity 3.169569e-07          TRUE
```

The number of proteins, can be found in the Data Processing file. Characteristic path length, network density, nodes, and edges calculated from Cytoscape built-in statistics.

Table S4

This data file is the same one that is generated by the data processing file. It is included in the same GitHub repo as this markdown file.

```
fullList <- read.csv("fullList.csv", header = TRUE, stringsAsFactors =
FALSE)
```

First get data into correct form. The `oorToNa()` function changes out-of-range genes to NAs that way numeric calculations can be performed.

```
oorToNa <- function(geneList) {
    geneList[geneList == "LOW"] <- NA
    geneList[geneList == "HIGH"] <- NA
    geneList[geneList == "<NA>"] <- NA

    return(geneList)
}

fullListNumeric <- oorToNa(fullList)
```

Now we can output the PSD proteins seen in table S4 (available online). The comment here indicates there are 76 proteins as reported in the paper.

```
subset(fullListNumeric, as.numeric(PSD) > -0.1 & as.numeric(PSD) < 0.1,
c("Gene", "PSD"))   # 76 proteins
```

Table S5

Output of the PSD proteins seen in table S5 (available online). The comments here indicate that there were either 159 or 166 proteins in each of the two directions for the PSD.

```
subset(fullListNumeric, as.numeric(PSD) < -0.1, c("Gene", "PSD"))  # 15
9 proteins

subset(fullListNumeric, as.numeric(PSD) > 0.1, c("Gene", "PSD"))   # 16
6 proteins
```

Table S6

Output of the out-of-range proteins in the PSD. Note that we went back and manually removed proteins where the fold-change calculation may have been unreliable - in cases where there was either high variability between samples or where both genes only had data for 1 out of 3 spectral counts for the replicates in the numerator or denominator.

```
as.vector(na.omit(fullList[fullList$PSD == "HIGH", "Gene"]))

##  [1] "Actr2"    "Aldh2"    "Aldh6a1"  "Anxa6"    "Apoe"     "Arpc5l"
##  [7] "Atp5j"    "Cadm4"    "Cct4"     "Clasp2"   "Col1a1"   "Dctn2"
## [13] "Ddost"    "Dync1li1" "Eef1d"    "Efhd2"    "Eif5a"    "Erp29"
## [19] "Fahd2"    "Glo1"     "Gpd1"     "Gpr158"   "Grin1"    "H2AFR"
## [25] "Hsd11b1"  "L1cam"    "Lap3"     "Lin7c"    "Macf1"    "Maoa"
## [31] "Mapk10"   "Marcks"   "Nckap1"   "Nf1"      "Ngef"     "Oat"
## [37] "Pcmt1"    "Pdzd2"    "Ppib"     "Prkca"    "Prkcb"    "Qdpr"
## [43] "Rph3a"    "Rplp0"    "Slc32a1"  "Slc9a3r1" "Suclg1"   "TSH2B"

as.vector(na.omit(fullList[fullList$PSD == "LOW", "Gene"]))

##  [1] "Acsf2"     "Aifm1"     "Ap2s1"     "Atad3"     "Atl1"
##  [6] "Atp1b3"    "Cct5"      "Copg1"     "Crebbp"    "Ctbp1"
```

```
## [11] "Cyb5r1"    "Dctn1"     "Dlg2"      "Dlg3"      "Dpysl5"
## [16] "Eef2"      "Eif4a2"    "Epb41l1"   "Gja1"      "Gphn"
## [21] "Hivep2"    "Hsd17b10"  "Hsph1"     "Idh2"      "Impact"
## [26] "Kidins220" "Kif2a"     "Map2"      "Map6"      "Mtnd4"
## [31] "Myo5a"     "Nap1l1"    "Ndufa11"   "P4hb"      "Park7"
## [36] "Pclo"      "Pde2a"     "Pgm1"      "Pik3cb"    "Plcb1"
## [41] "Plec"      "Ppp1r9b"   "Pygm"      "Rpn2"      "Rtn3"
## [46] "Rtn4"      "Rufy3"     "Ryr1"      "Slc30a3"   "Slc8a2"
## [51] "Snd1"      "Syt12"     "Thop1"     "Tkt"       "Tmem39a"
## [56] "Vat1"      "Wdr1"      "Zc3h18"
```

Table S7

See the figure 6 markdown file to generate these lists.

Table S8

See the figure 6 markdown file to generate these lists.

All other figures generated through either Cytoscape or DAVID or the output of the mass spectrometry data analysis.

Appendix D: Code for Chapter 4

UBUNTU INSTALL SHELL SCRIPT

This section contains the shell commands followed on a fresh Ubuntu install to add all software used during RNAseq analysis. This is provided for reproducibility. Note that the shell commands here do not form a complete shell script so they should not be run directly. Please modify as needed.

```
# Install Zshell
sudo apt-get install zsh
sudo apt-get install git-core
sudo wget https://github.com/robbyrussell/oh-my-zsh/raw/master/tools/in
stall.sh -O - | zsh
chsh -s `which zsh`
sudo shutdown -r 0

# Install R Base
sudo echo "deb http://cran.rstudio.com/bin/linux/ubuntu xenial/" | sudo
tee -a /etc/apt/sources.list
gpg --keyserver keyserver.ubuntu.com --recv-key E084DAB9
gpg -a --export E084DAB9 | sudo apt-key add -
sudo apt-get update
sudo apt-get install r-base r-base-dev

# Install R Studio 0.99.903
sudo apt-get install gdebi-core
wget https://download1.rstudio.org/rstudio-0.99.903-amd64.deb
sudo gdebi -n rstudio-0.99.903-amd64.deb
rm rstudio-0.99.903-amd64.deb

# Install dropbox
sudo apt-get install nautilus-dropbox

# Run dropbox to initiate download

# Install Java/Oracle 8
sudo apt-get update
sudo apt-get install default-jre
sudo apt-get install default-jdk
sudo add-apt-repository ppa:webupd8team/java
sudo apt-get update
sudo apt-get install oracle-java8-installer
```

```
# Install FastQC 0.11.5
cd /usr/src
sudo curl -O http://www.bioinformatics.babraham.ac.uk/projects/fastqc/f
astqc_v0.11.5.zip
sudo unzip fastqc_v0.11.5.zip
sudo ln -sf /usr/src/FastQC/fastqc /usr/bin/fastqc
sudo chmod +x /usr/bin/fastqc

# Install pip
sudo apt-get update
sudo apt-get upgrade
sudo apt-get install python-pip

# Install python 2.7.10
sudo apt-get install build-essential checkinstall
sudo apt-get install libreadline-gplv2-dev libncursesw5-dev libssl-dev
libsqlite3-dev tk-dev libgdbm-dev libc6-dev libbz2-dev
cd /usr/src
sudo wget https://www.python.org/ftp/python/2.7.10/Python-2.7.10.tgz
sudo tar xzf Python-2.7.10.tgz
cd Python-2.7.10
sudo ./configure
sudo make altinstall
sudo apt-get install python-dev

# Install cutadapt 1.11
sudo -H pip install cutadapt

# Install Trim Galore 0.4.1
cd /usr/src
sudo curl -O http://www.bioinformatics.babraham.ac.uk/projects/trim_gal
ore/trim_galore_v0.4.1.zip
sudo unzip trim_galore_v0.4.1.zipsudo ln -sf /usr/src/trim_galore_zip/t
rim_galore /usr/bin/trim_galore

# Install Tophat 2.1.1
cd /usr/src
sudo curl -O https://ccb.jhu.edu/software/tophat/downloads/tophat-2.1.1
.Linux_x86_64.tar.gz
sudo tar xzvf tophat-2.1.1.Linux_x86_64.tar.gz

# Install Bowtie 2.2.9
cd /usr/src
sudo curl -OL http://sourceforge.net/projects/bowtie-bio/files/bowtie2/
2.2.9/bowtie2-2.2.9-linux-x86_64.zip
```

```
sudo unzip bowtie2-2.2.9-linux-x86_64.zip

# Install samtools 0.1.19
cd /usr/src
sudo curl -OL https://sourceforge.net/projects/samtools/files/samtools/
0.1.19/samtools-0.1.19.tar.bz2
sudo tar jxvf samtools-0.1.19.tar.bz2
sudo chown $USER -R /usr/src/samtools-0.1.19
cd samtools-0.1.19
sudo make

# Install RSeQC 2.6.4
sudo -H pip install RSeQC

# Add the following paths to the zshell profile:
export PATH=/usr/src/tophat-2.1.1.Linux_x86_64:$PATH
export PATH=/usr/src/bowtie2-2.2.9:$PATH
export PATH=/usr/src/samtools-0.1.19:$PATH
export PATH=/usr/src/piranha-1.2.1:$PATH

## The following contains all shell commands for NGS analysis tools
# Install Piranha 1.2.1
cd /usr/src
sudo curl -O http://smithlabresearch.org/downloads/piranha-1.2.1.tar.gz
sudo tar xzvf piranha-1.2.1.tar.gz
cd piranha-1.2.1
sudo ./configure --with-bam_tools_headers="/usr/src/BAMTools/include/"
\
      --with-bam_tools_library="/usr/src/BAMTools/lib/"
sudo make all
sudo make install

# Install CMake 3.6.2
cd /usr/src
sudo curl -O https://cmake.org/files/v3.6/cmake-3.6.2.tar.gz
sudo tar xzvf cmake-3.6.2.tar.gz
cd cmake-3.6.2
sudo chmod u+x bootstrap
sudo ./bootstrap
sudo make
sudo make install

# Install BAMtools 2.4.1
cd /usr/src
sudo git clone git://github.com/pezmaster31/bamtools.git
cd bamtools
```

```
sudo mkdir build
cd build
cd cmake ..
sudo make

# Install GSL-2.2.1
cd /usr/src
sudo curl -O http://gnu.mirrors.pair.com/gnu/gsl/gsl-2.2.1.tar.gz
sudo tar xzvf gsl-2.2.1.tar.gz
sudo ./configure
sudo make
sudo make install
sudo apt-get install libgsl0-dev

# Install BEDtools 2.25.0-1
sudo apt-get install bedtools

# Prepare to install R/Bioconductor packages
sudo chown $USER -R /usr/lib/R
sudo chown $USER -R /usr/share/R
sudo apt-get update
sudo apt-get install r-base-html r-base-latex r-cran-rgtk2 r-cran-rggob
i r-cran-rgl libxml2-dev libcurl4-gnutls-dev

# Install R packages
install.packages("R.utils")
install.packages("ggplot2")
install.packages("RColorBrewer")
install.packages("pheatmap")
install.packages("reshape2")

# Install Bioconductor and related packages
source("http://bioconductor.org/biocLite.R")
biocLite("org.Mm.eg.db")
biocLite("DESeq2")
biocLite("Rsubread")
biocLite("genefilter")
biocLite("edgeR")
biocLite("AnnotationDbi")
biocLite("RIPSeeker")
biocLite("biomaRt")
biocLite("ChIPpeakAnno")
```

QC, TRIMMING, ALIGNMENT, AND COUNTS

There are two different sets of technical replicates reflecting the runs from July 2016 (two replicates) and August 2016 (two replicates). QC and trimming was performed separately on each set of runs. Reads from the runs were combined together for analysis after alignment.

Load packages

```
library(R.utils)
library(Rsubread)
```

Quality Control

Command: `fastqc`

QC was performed with FastQC **(version 0.11.5)** in the terminal. Then, all FASTQ files were moved to a new directory.

```
system("fastqc *.gz")
dir.create("fastqc")
system("mv *.zip *.html fastqc")
```

The QC files show high-quality Q-scores across all sequences. There is minor adaptor contamination. Repetitive sequences are generally below 1% of all sequences in a given library and consist primarily of ribosomal RNA or adaptors. The replicates from the Ro knockout show low library complexity and unusual GC content (compared to the other libraries). This is likely because of the small amount of starting material. We were only able to obtain one replicate for the Ro knockout because the other replicates had too little material to sequence. Overall, the QC shows that the data is of high quality.

Trimming

Command: `trim_galore --fastqc -o trimmed_fastqc *.gz`

Adaptor/quality trimming was performed with **Trim Galore (version 0.4.1)** using **Python 2.7.11**. Trim Galore is wrapped around **cutadapt 1.10** and has many known adaptors pre-loaded so they do not need to be specified in advance. The basic options for trim galore do not perform any aggressive trimming and in general, a high amount of trimming was not needed.

```
dir.create("trimmed_fastqc")
system("trim_galore --fastqc -o trimmed_fastqc *.gz")
```

Alignment

Command: `tophat2 library-type fr-firststrand -G [GTF] -p 4 -o [library][index][files]`

For alignment, we first created a metatable to help organize all the different libraries which aids in downstream processing steps. The following code chunk also creates a new alignment directory:

```
metaTable <- data.frame(LibraryName = c("DMSO1_T1", "DMSO1_T2", "DMSO2_
T1", "DMSO2_T2",
                                        "RO_KO2_T1", "RO_KO2_T2",
                                        "RO_RAP_KO1_T1", "RO_RAP_KO1_T2
", "RO_RAP_KO2_T1", "RO_RAP_KO2_T2", "RO_RAP_KO3_T1", "RO_RAP_KO3_T2",
                                        "RO_RAP_WT1_T1", "RO_RAP_WT1_T2
", "RO_RAP_WT2_T1", "RO_RAP_WT2_T2", "RO_RAP_WT3_T1", "RO_RAP_WT3_T2",
                                        "RO_WT1_T1", "RO_WT1_T2", "RO_W
T2_T1", "RO_WT2_T2", "RO_WT3_T1", "RO_WT3_T2",
                                        "SAL_KO1_T1", "SAL_KO1_T2", "SA
L_KO2_T1", "SAL_KO2_T2", "SAL_KO3_T1", "SAL_KO3_T2",
                                        "SAL_WT1_T1", "SAL_WT1_T2", "SA
L_WT2_T1", "SAL_WT2_T2", "SAL_WT3_T1", "SAL_WT3_T2",
                                        "TOT1_T1", "TOT1_T2", "TOT2_T1"
, "TOT2_T2", "TOT3_T1", "TOT3_T2"),
                        LibraryLayout = rep("SINGLE", 42),
```

```
                        fastq = grep("fq.gz", list.files(getwd()), valu
e = T),
                        condition = c(rep("CTL", 4), rep("KO", 8), rep(
"WT", 12), rep("KO", 6), rep("WT", 6), rep("TOT", 6)),
                        technical = rep(1:2, 21))

metaTable$LibraryName <- as.factor(metaTable$LibraryName)

dir.create("alignment")
setwd("alignment")
```

Gene model annotations (mm9) were downloaded from Ensembl.org and pre-built reference indices were downloaded from the Illumina website.

```
### Gene model annotation
URL1 <- "ftp://ftp.ensembl.org/pub/release-84/gtf/mus_musculus/Mus_musc
ulus.GRCm38.84.gtf.gz"
download.file(URL1, "Mus_musculus.GRCm38.84.gtf.gz")
gunzip("Mus_musculus.GRCm38.84.gtf.gz")

### Pre-built reference indices
URL2 <- "ftp://igenome:G3nom3s4u@ussd-ftp.illumina.com/Mus_musculus/Ens
embl/GRCm38/Mus_musculus_Ensembl_GRCm38.tar.gz"
download.file(URL2, "Mus_musculus_Ensembl_GRCm38.tar.gz")
gunzip("Mus_musculus_Ensembl_GRCm38.tar.gz")
untar("Mus_musculus_Ensembl_GRCm38.tar")
```

To align reads to genome, we used **Tophat 2.1.1**, **Bowtie 2.2.9**, and **samtools 0.1.19**. The gtf and index variables are strings leading to the directory where these files are stored. The library type was specified as fr-firststrand in accordance with the library prep protocol.

```
### Move all trimmed FASTQ data files to the alignment directory
setwd("..")
system("mv *.gz alignment")
setwd("alignment")

gtf <- "/media/sanjeev/MyPassport/Sanjeev/fmrp_ip_analysis/fastq_data/t
rimmed_fastqc/alignment/Mus_musculus.GRCm38.84.gtf"
index <- "/media/sanjeev/MyPassport/Sanjeev/fmrp_ip_analysis/fastq_data
/trimmed_fastqc/alignment/Mus_musculus/Ensembl/GRCm38/Sequence/Bowtie2I
ndex/genome"
```

```
### Create a vector that stores the tophat terminal commands for each f
ile
tophat <- with(metaTable, paste("tophat2 --library-type fr-firststrand
-G ", gtf, "-p 4 -o", LibraryName, index, fastq))

### This loop runs all the tophat commands in the terminal
for(i in 1:length(tophat)) {
  system(tophat[i])
}
```

We created a table from the alignment results file found in each Tophat alignment file. These data can be found in the full_alignment_summary.csv file.

```
output = data.frame()

for(i in 1:length(metaTable$LibraryName)) {
  dat <- read.table(paste(metaTable[i,1], "/align_summary.txt", sep = "
"), header = FALSE, sep = "\t", fill = TRUE, stringsAsFactors = FALSE)
  input <- substr(dat[2,1],25,nchar(dat[2,1]))
  mapped <- substr(dat[3,1],25,nchar(dat[3,1]))
  multiple <- substr(dat[4,1],26,nchar(dat[4,1]))
  rate <- substr(dat[5,1], 1, nchar(dat[5,1])-1)

  summary <- data.frame(metaTable[i,1], input, mapped, multiple, rate)

  output <- rbind(output, summary)
}

names(output) <- c("Library", "Input_Reads", "Mapped_Reads", "Multiple_
Alignments", "Mapping_Rate")

write.csv(output, "full_alignment_summary.csv")   # Save alignment stat
s to working dir
```

The alignment data shows that most libraries had relative strong alignment percentages (greater than 75%). The only exception to this is the Ro KO replicate 2. As mentioned previously, there was very low starting material for this library and low library complexity. While some of this can be explaining through ribosomal sequence contamination, there are

also further sequences in this sample that may be useful in differential gene expression analysis. Thus, we have opted not to remove this library.

Calculate counts from BAM files

We used **Rsubread 1.20.6** to calculate counts. The strandSpecific = 2 option was chosen in accordance with the library prep protocol. We used the countMultiMappingReads = TRUE option specifically for the purpose of counting ambiguously mapped reads. At this point, all BAM files from both runs (July 2016 and August 2016) were combined into one directory so that the the final output includes counts for four technical replicates.

```
### Create an extry column for counts and a new vector that contains th
e directory for all the BAM files from Tophat output
metaTable$Counts <- paste(metaTable$LibraryName, "count", sep = ".")
alignmentFilesDir <- file.path(getwd(), paste0(metaTable$LibraryName, "
/", "accepted_hits.bam"))

path <- "/media/sanjeev/MyPassport/Sanjeev/fmrp_ip_analysis/fastq_data/
trimmed_fastqc/alignment"
gtfDir <- file.path(path, "Mus_musculus.GRCm38.84.gtf")

### Run Rsubread on all BAM files
for(i in 1:length(metaTable$LibraryName)) {
  assign(paste0(metaTable[i,1],"_Count"), featureCounts(alignmentFilesD
ir[i],
                                                        annot.ext = gtf
Dir,
                                                        isGTFAnnotation
File = T,
                                                        isPairedEnd = F
ALSE,
                                                        countMultiMappi
ngReads = TRUE,
                                                        strandSpecific
= 2,
                                                        nthreads = 4))
}

### Bind columns together into one dataframe and export
```

```
listObjects <- lapply(grep("*Count",ls(), value = T), get)
countData <- as.data.frame(lapply(listObjects, function(x) x$counts))
names(countData) <- metaTable$LibraryName
write.csv(countData, "countData.csv")
```

Session info

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## loaded via a namespace (and not attached):
##  [1] backports_1.0.4 magrittr_1.5    rprojroot_1.1   tools_3.3.2
##  [5] htmltools_0.3.5 yaml_2.1.14     Rcpp_0.12.8     stringi_1.1.2
##  [9] rmarkdown_1.2   knitr_1.15.1    stringr_1.1.0   digest_0.6.10
## [13] evaluate_0.10
```

POST-PROCESSING QC

Here we perform some exploratory data analysis in order to confirm the quality of our data.

This includes PCA, hierarchical clustering between the different groups, and correlation of counts among the replicates.

Load packages

```
library(DESeq2)
library(magrittr)
library(ggplot2)
library(gridExtra)
library(pheatmap)
library(RColorBrewer)
```

Load data and process

The count data from the different sequencing runs are loaded and combined together into one unit. Next, a metatable is created for easier downstream processing. Then, technical replicates are collapsed by DESeq2 to create the final dataset used for the QC analysis. We apply the regularized log transformation to the data so that it is more Gaussian, a requirement for the QC methods.

```
### Load count data
countData1 <- read.csv("countData_july.csv", header = TRUE)
rownames(countData1) <- countData1[ ,1]
countData1[ ,1] <- NULL

countData2 <- read.csv("countData_aug.csv", header = TRUE)
countData2[ ,1] <- NULL
names(countData2) <- gsub("_T1", "_T3", names(countData2))
names(countData2) <- gsub("_T2", "_T4", names(countData2))

countData <- cbind(countData1, countData2) %>% as.data.frame()

### Create metatable
metaTable <- data.frame(LibraryName = names(countData),
                        technical = c(rep(1:2, 21), rep(3:4, 21)))

metaTable$LibraryName <- as.factor(metaTable$LibraryName)
metaTable$technical <- as.factor(metaTable$technical)
metaTable$Counts <- paste(metaTable$LibraryName, "count", sep = ".")

rownames(metaTable) <- metaTable[ ,1]   # Set first column to row names
metaTable[ ,1] <- NULL

### Collapse technical replicates
dd <- DESeqDataSetFromMatrix(countData = countData,
                             colData = metaTable,
                             design = ~ technical)

dd$technical <- gsub("_T1|_T2|_T3|_T4", "", row.names(metaTable))
ddRep <- collapseReplicates(dd, groupby = dd$technical)
ddRep$technical <- as.factor(ddRep$technical)

### rlog transformation of data
```

```
rd <- rlog(ddRep)
ddRep <- estimateSizeFactors(ddRep)
rdCounts <- assay(rd) %>% as.data.frame()
```

Count correlations

A simple way to visualize the precision of our independent biological replicates is simply by graphing their counts against each other. We expect a high correlation between all different groups. The data shows a high amount of correlation between log-transformed counts among the different groups.

```
### Saline WT scatterplots
p2 <- ggplot(rdCounts, aes(SAL_WT1, SAL_WT2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p3 <- ggplot(rdCounts, aes(SAL_WT2, SAL_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p4 <- ggplot(rdCounts, aes(SAL_WT1, SAL_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Saline KO scatterplots
p5 <- ggplot(rdCounts, aes(SAL_KO1, SAL_KO2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p6 <- ggplot(rdCounts, aes(SAL_KO2, SAL_KO3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p7 <- ggplot(rdCounts, aes(SAL_KO1, SAL_KO3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Ro WT scatterplots
p8 <- ggplot(rdCounts, aes(RO_WT1, RO_WT2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()
```

```
p9 <- ggplot(rdCounts, aes(RO_WT2, RO_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p10 <- ggplot(rdCounts, aes(RO_WT1, RO_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Ro KO histogram
p11 <- ggplot(rdCounts, aes(RO_KO2, RO_KO2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Ro Rap WT scatterplots
p12 <- ggplot(rdCounts, aes(RO_RAP_WT1, RO_RAP_WT2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p13 <- ggplot(rdCounts, aes(RO_RAP_WT2, RO_RAP_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p14 <- ggplot(rdCounts, aes(RO_RAP_WT1, RO_RAP_WT3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Ro Rap KO scatterplots
p15 <- ggplot(rdCounts, aes(RO_RAP_KO1, RO_RAP_KO2)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p16 <- ggplot(rdCounts, aes(RO_RAP_KO2, RO_RAP_KO3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

p17 <- ggplot(rdCounts, aes(RO_RAP_KO1, RO_RAP_KO3)) +
  geom_point(size = 0.5, alpha = 0.3) +
  theme_bw()

### Fully arranged (766 x 673)
grid.arrange(p2, p3, p4, p5,
             p6, p7, p8, p9,
             p10, p11, p12, p13,
             p14, p15, p16, p17,
             ncol = 4,
             nrow = 4)
```

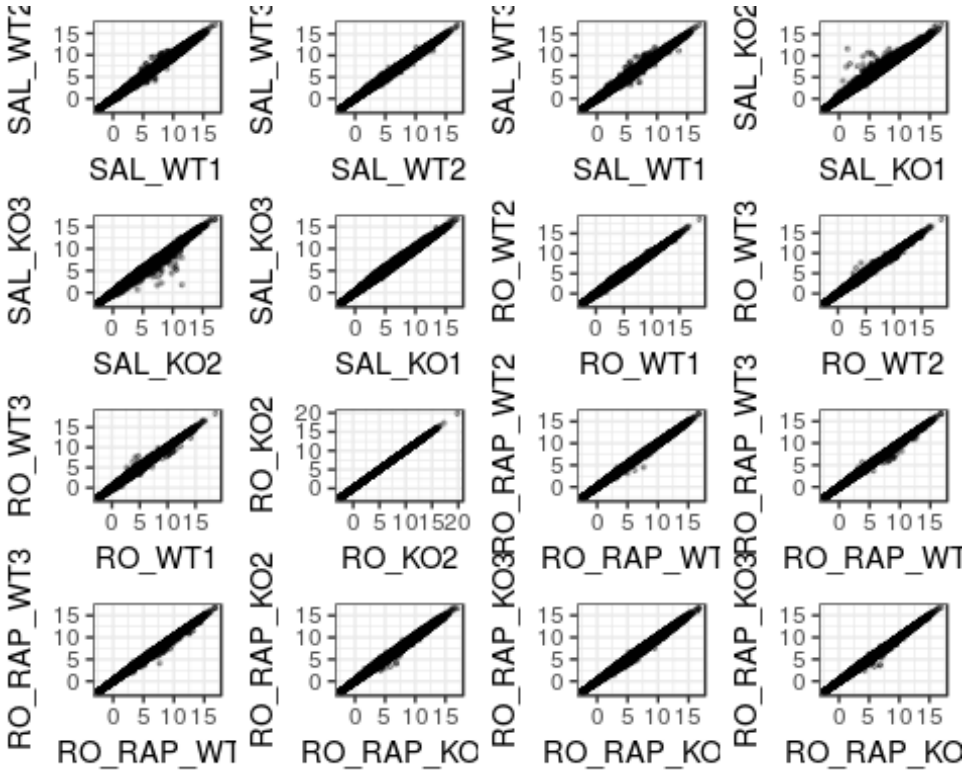


Figure D.1: Count correlations between RIP-seq libraries

Principal component analysis (PCA)

Principal component analysis is another means of visualizing the similarity in count measurements between the different groups. The adjusted axes allow us to see whether or not the different groups cluster together. The data shows that each replicate for WT and KO cluster together due to similarity in variance. The only notable exception is the saline KO replicate 2 which is slightly off compared to the others. Similarly, the Ro KO replicate 2 is also quite far off from the others. This is not completely unexpected however, since we did note a poor alignment rate and unusual QC from the FastQC files.

```
# Arrange variables for plotting with ggplot2
rdNoTotal <- rd[ ,rd$technical %in% colData(rd)[3:18, 1]]
pcaDatNoTotal <- plotPCA(rdNoTotal, intgroup = "technical", returnData
= TRUE)
percentVar <- round(100 * attr(pcaDatNoTotal, "percentVar"))

ggplot(pcaDatNoTotal, aes(PC1, PC2, color = technical, label = name)) +
  geom_point(size = 3) +
  geom_text(aes(label = name), size = 3, hjust = 0.3, vjust = -0.5) +
  xlab(paste0("PC1: ",percentVar[1],"% variance")) +
  ylab(paste0("PC2: ",percentVar[2],"% variance")) +
  geom_hline(aes(yintercept = 0)) +
  geom_vline(aes(xintercept = 0)) +
  scale_color_manual(values = c(rep("darkolivegreen3", 1), rep("cadetbl
ue3", 3), rep("dodgerblue3", 3), rep("chartreuse4", 3), rep("lightcoral
", 3), rep("firebrick", 3))) +
  theme_bw() +
  theme(legend.position = "none")
```

Euclidean distance matrix

Finally, we used a clustered heatmap to visualize the similarity between the different samples. Overall, many groups clustered together though some replicates, particularly the Saline KO replicate 2, clustered better with other libraries than it did its own replicates. It is contained in a branch on its own which is a parent branch to two other libraries Ro+Rapa KO replicate 2 and Saline KO replicate 3 which are also off on their own. In general, since the KO libraries are constructed from null-antibody tissue we expect their to be some level of background binding that will be inconsistant among the different samples. Since the wild-type material clusters together properly, we have kept all libraries in going forward.

```
sampleDists <- dist(t(assay(rdNoTotal)))

sampleDistMatrix <- as.matrix(sampleDists)
rownames(sampleDistMatrix) <- colnames(rdNoTotal)
colnames(sampleDistMatrix) <- NULL
colors <- colorRampPalette(rev(brewer.pal(9, "Blues")))(255)
```

```
pheatmap(sampleDistMatrix,
         clustering_distance_rows = sampleDists,
         clustering_distance_cols = sampleDists,
         col = colors)
```

Session info

```
sessionInfo()
```

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] parallel  stats4    stats     graphics  grDevices utils     data
sets
## [8] methods   base
##
## other attached packages:
##  [1] RColorBrewer_1.1-2         pheatmap_1.0.8
##  [3] gridExtra_2.2.1            ggplot2_2.2.0
##  [5] magrittr_1.5               DESeq2_1.14.0
##  [7] SummarizedExperiment_1.4.0 Biobase_2.34.0
##  [9] GenomicRanges_1.26.1       GenomeInfoDb_1.10.1
## [11] IRanges_2.8.1              S4Vectors_0.12.0
## [13] BiocGenerics_0.20.0
##
## loaded via a namespace (and not attached):
##  [1] genefilter_1.56.0    locfit_1.5-9.1       splines_3.3.2
##  [4] lattice_0.20-34      colorspace_1.3-1     htmltools_0.3.5
##  [7] yaml_2.1.14          survival_2.40-1      XML_3.98-1.5
## [10] foreign_0.8-67       DBI_0.5-1            BiocParallel_1.8.1
## [13] plyr_1.8.4           stringr_1.1.0        zlibbioc_1.20.0
## [16] munsell_0.4.3        gtable_0.2.0         evaluate_0.10
## [19] memoise_1.0.0        labeling_0.3         latticeExtra_0.6-28
## [22] knitr_1.15.1         geneplotter_1.52.0   AnnotationDbi_1.36.0
## [25] htmlTable_1.7        Rcpp_0.12.8          acepack_1.4.1
## [28] xtable_1.8-2         scales_0.4.1         backports_1.0.4
## [31] Hmisc_4.0-0          annotate_1.52.0      XVector_0.14.0
```

```
## [34] digest_0.6.10        stringi_1.1.2        grid_3.3.2
## [37] rprojroot_1.1        tools_3.3.2          bitops_1.0-6
## [40] lazyeval_0.2.0       RCurl_1.95-4.8       tibble_1.2
## [43] RSQLite_1.1          Formula_1.2-1        cluster_2.0.5
## [46] Matrix_1.2-7.1       data.table_1.9.8     assertthat_0.1
## [49] rmarkdown_1.2        rpart_4.1-10         nnet_7.3-12
```

COUNT NORMALIZATION

Here we annotate our data in official gene symbol format and apply DESeq2 factor size normalization across all our libraries. We also perform filtering and some visualization of the count data before the final export.

Load packages

```
library(DESeq2)
library(AnnotationDbi)
library(org.Mm.eg.db)
library(HTSFilter)
library(magrittr)
library(ggplot2)
library(reshape2)
library(ggthemes)
```

Functions

The squash() function simply scales any data between 0-100 for easier visualization and comparison.

```
squash <- function(x) {
  output <- ((x - min(x)) / (max(x) - min(x))) * 100
  return(output)
}
```

The percentChange() function calculates the percent a measure changes from its baseline.

```
percentChange <- function(new, old) {
  return(((new - old) / abs(old)) * 100)
}
```

Count filtering

We used **HTSFilter 1.14.1** to filter the data using the standard settings.

```
countData <- read.csv(file = "combinedCounts.csv", header = TRUE)
rownames(countData) <- countData[ ,1]
countData[ ,1] <- NULL

countData$RO_KO1 <- countData$RO_KO2
countData$RO_KO3 <- countData$RO_KO2

cdMain <- subset(countData, select = c(16:18, 13:15, 10:12, 3, 22, 23,
7:9, 4:6))
conds <- sapply(names(cdMain), function(x) substr(x, 1, nchar(x)-1)) %>
% unname

filter <- HTSFilter(x = cdMain,
                    conds = conds,
                    normalization = "none")
```

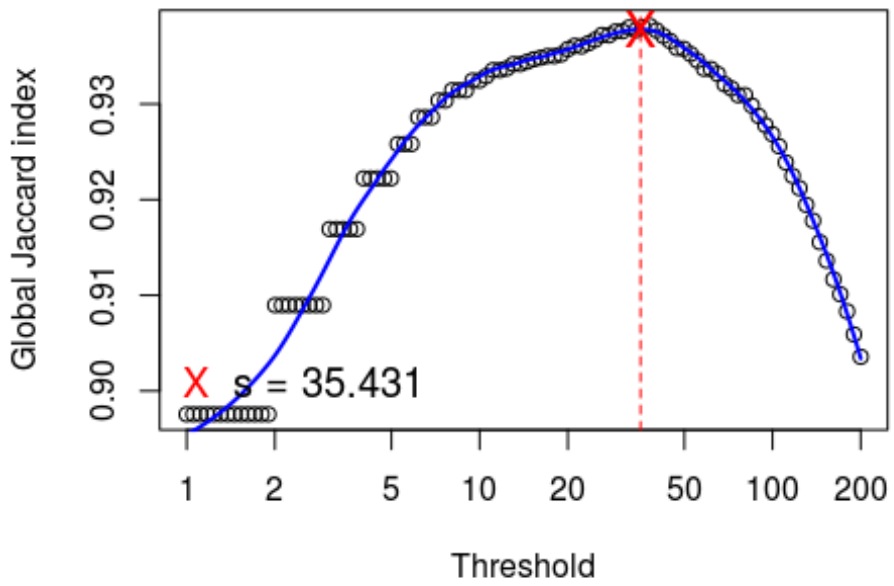


Figure D.2: Global Jaccard index for low count filter

```
cdFilter <- as.data.frame(filter$filteredData)
cdFilter[ ,c(11, 12)] <- NULL   # Remove duplicated Ro KO columns
```

Annotation (official gene symbol)

Annotation weas performed with the Bioconductor packages **org.Mm.eg.db** and **AnnotationDbi**. Any sequences that did not map to a gene qualifier were filtered out of the dataset.

```
cdFilter$symbol <- mapIds(org.Mm.eg.db,
                      keys = row.names(cdFilter),
                      column = "SYMBOL",
                      keytype = "ENSEMBL",
                      multiVals = "first")

### Remove rows with no gene mapping
cdFilter <- subset(cdFilter, select = c(17, 1:16))
cdFilter <- cdFilter[complete.cases(cdFilter), ]
rownames(cdFilter) <- make.names(cdFilter$symbol, unique = TRUE)
cdFilter[ ,1] <- NULL
```

Count normalization (DESeq2 factor size)

First, we create a metatable to organize the data. Next, the data is loaded into **DESeq2**. Since DESeq2 scaling applies across the entire library, it inflates the counts from our KO libraries. To fix this, we scale back the KO counts using ratios between WT and KO counts obtained from the BioAnalyzer results prior to sequencing. Then, all counts were collapsed by averaging across rows.

```
### Create metatable
metaTable <- data.frame(LibraryName = names(cdFilter),
                        LibraryLayout = rep("SINGLE", 16),
                        condition = c(rep("WT", 3), rep("KO", 3), rep("
WT", 3), "KO", rep("WT", 3), rep("KO", 3)))
rownames(metaTable) <- metaTable[ ,1]   # Set first column to row names
metaTable[ ,1] <- NULL

### Get normalized counts
```

```
dd <- DESeqDataSetFromMatrix(countData = cdFilter,
                             colData = metaTable,
                             design = ~ 1)
dd <- estimateSizeFactors(dd)
nc <- as.data.frame(counts(dd, normalize = TRUE))

### Scale by BioAnalyzer values
BA <- read.csv("BA_values.csv", header = TRUE, stringsAsFactors = FALSE
)
BA$mean <- rowMeans(BA[ ,2:4], na.rm = TRUE)

scale <- c(RO = BA$mean[2] / BA$mean[5],
           RO_RAP = BA$mean[3] / BA$mean[4],
           SAL = BA$mean[6] / BA$mean[7])

cdFilterScaled <- nc

cdFilterScaled[ ,10] <- cdFilterScaled[ ,10] / (scale[1] + 1)         #
Scale Ro
cdFilterScaled[ ,14:16] <- cdFilterScaled[ ,14:16] / (scale[2] + 1)   #
Scale Ro Rap
cdFilterScaled[ ,4:6] <- cdFilterScaled[ ,4:6] / (scale[3] + 1)       #
Scale Sal

### Average across rows
nscMean <- data.frame(row.names = row.names(cdFilterScaled),
                      SAL_WT = rowMeans(cdFilterScaled[ ,1:3]),
                      SAL_KO = rowMeans(cdFilterScaled[ ,4:6]),
                      RO_WT = rowMeans(cdFilterScaled[ ,7:9]),
                      RO_KO = cdFilterScaled[ ,10],
                      RO_RAP_WT = rowMeans(cdFilterScaled[ ,11:13]),
                      RO_RAP_KO = rowMeans(cdFilterScaled[ ,14:16])) %>%
squash()
```

Visualize normalized counts

For count visualization, we first plotted the WT and KO counts be each treatment

```
### Create variables for plotting with ggplot2
nscMean.m <- melt(nscMean)
nscMean.m$Genotype <- c(rep("WT", 14543), rep("KO", 14543), rep("WT", 1
4543), rep("KO", 14543), rep("WT", 14543), rep("KO", 14543))
nscMean.m$variable <- c(rep("Saline", 29086), rep("Ro-25-6981", 29086),
rep("Ro-25-6981 + Rapamycin", 29086))
nscMean.m$varOrder <- factor(nscMean.m$variable, levels = c("Saline", "
```

```
Ro-25-6981", "Ro-25-6981 + Rapamycin"))

### Plot WT v. KO by treatment
ggplot(nscMean.m, aes(x = Genotype, y = value, color = Genotype)) +
  geom_point(position = position_jitter(width = 0.2, height = 0.2)) +
  facet_grid(~ varOrder) +
  xlab("") +
  ylab("Scaled Counts") +
  theme_few() +
  theme(axis.ticks.x = element_blank(),
        axis.text.x = element_blank())
```

Next, we plotted WT and KO counts with known FMRP targets and non-targets (Darnell et al., 2011)

```
### Load targets and non-targets
targets <- c("Gabbr1", "Gabbr2","Nisch", "Prrc2a", "Kif1a", "Grin2a", "
Adgrb2", "Cyfip2", "Map1b", "Bsn", "Adcy1", "Pde2a", "Lingo1", "Grin2b"
)
nontargets <- c("Pabpc1", "Tmem65", "Hprt1", "St8sia3", "Sae1", "Glrb",
"Gria2", "Tcerg1", "Eif3a", "Slc35f1", "Atp6ap2", "Vldlr")

### Prepare variables for plotting
nscRawTargets <- nscMean[row.names(nscMean) %in% c(targets, nontargets)
, ] %>% squash()
nscRawTargets$targets <- ifelse(row.names(nscRawTargets) %in% targets,
"Target", "Non-target")
nscRawTargets.m <- melt(nscRawTargets)
nscRawTargets.m$Genotype <- c(rep("WT", 25), rep("KO", 25), rep("WT", 2
5), rep("KO", 25), rep("WT", 25), rep("KO", 25))
nscRawTargets.m$variable <- c(rep("Saline", 50), rep("Ro-25-6981", 50),
rep("Ro-25-6981 + Rapamycin", 50))
nscRawTargets.m$group <- rep(1:25, 6)
nscRawTargets.m$varOrder <- factor(nscRawTargets.m$variable, levels = c
("Saline", "Ro-25-6981", "Ro-25-6981 + Rapamycin"))
nscRawTargets.m$GenotypeOrder <- factor(nscRawTargets.m$Genotype, level
s = c("WT", "KO"))

### Plot WT v. KO for FMRP targets and non-targets
ggplot(nscRawTargets.m, aes(x = GenotypeOrder, y = value, group = group
, color = targets)) +
  geom_point() +
  geom_line() +
```

```
  facet_grid(~ varOrder) +
  xlab("") +
  ylab("Scaled Counts") +
  theme_few() +
  scale_color_discrete(name = "FMRP Targets")
```

Finally, for the percent changes between targets and non-targets reported in the paper, we simply calculated the percent changes across all the dataframe for the mean normalized (scaled) counts.

```
new <- nscRawTargets[nscRawTargets$targets == "Target", 1:6] %>% colMeans()
old <- nscRawTargets[nscRawTargets$targets == "Non-target", 1:6] %>% colMeans()

percentChange(new[1], old[1]) %>% round(2)

## SAL_WT
##  333.9

percentChange(new[3], old[3]) %>% round(2)

##  RO_WT
## 369.29

percentChange(new[5], old[5]) %>% round(2)

## RO_RAP_WT
##    365.35
```

Session info

```
sessionInfo()

## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
```

```
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] parallel  stats4    stats     graphics  grDevices utils     data
sets
## [8] methods   base
##
## other attached packages:
##  [1] ggthemes_3.3.0             reshape2_1.4.2
##  [3] ggplot2_2.2.0              magrittr_1.5
##  [5] HTSFilter_1.14.1           org.Mm.eg.db_3.4.0
##  [7] AnnotationDbi_1.36.0       DESeq2_1.14.0
##  [9] SummarizedExperiment_1.4.0 Biobase_2.34.0
## [11] GenomicRanges_1.26.1       GenomeInfoDb_1.10.1
## [13] IRanges_2.8.1              S4Vectors_0.12.0
## [15] BiocGenerics_0.20.0
##
## loaded via a namespace (and not attached):
##  [1] genefilter_1.56.0   locfit_1.5-9.1      splines_3.3.2
##  [4] lattice_0.20-34     colorspace_1.3-1    htmltools_0.3.5
##  [7] yaml_2.1.14         survival_2.40-1     XML_3.98-1.5
## [10] foreign_0.8-67      DBI_0.5-1           BiocParallel_1.8.1
## [13] RColorBrewer_1.1-2  plyr_1.8.4          stringr_1.1.0
## [16] zlibbioc_1.20.0     munsell_0.4.3       gtable_0.2.0
## [19] evaluate_0.10       memoise_1.0.0       labeling_0.3
## [22] latticeExtra_0.6-28 knitr_1.15.1        geneplotter_1.52.0
## [25] htmlTable_1.7       Rcpp_0.12.8         edgeR_3.16.4
## [28] acepack_1.4.1       xtable_1.8-2        scales_0.4.1
## [31] backports_1.0.4     limma_3.30.5        Hmisc_4.0-0
## [34] annotate_1.52.0     XVector_0.14.0      gridExtra_2.2.1
## [37] digest_0.6.10       DESeq_1.26.0        stringi_1.1.2
## [40] grid_3.3.2          rprojroot_1.1       tools_3.3.2
## [43] bitops_1.0-6        lazyeval_0.2.0      RCurl_1.95-4.8
## [46] tibble_1.2          RSQLite_1.1         Formula_1.2-1
## [49] cluster_2.0.5       Matrix_1.2-7.1      data.table_1.9.8
## [52] assertthat_0.1      rmarkdown_1.2       rpart_4.1-10
## [55] nnet_7.3-12
```

DATA FILTERING

This section contains all the code needed to generate the cutoff and filtration used for the knockout cutoff. Additionally, the filtration from Limma is also applied to the data.

Load packages

```
library(ggplot2)
library(reshape2)
library(biomaRt)
library(magrittr)
library(xlsx)
library(mgu74a.db)
library(AnnotationDbi)
library(org.Hs.eg.db)
library(seqinr)
library(grid)
library(gridExtra)
library(edgeR)
library(limma)
library(Ckmeans.1d.dp)
library(pheatmap)
library(RColorBrewer)
```

Functions

The cutoff() function is designed to create and above and below cutoff line using the consensus data sets. For every row in the data, it indicates the percentage of genes that are now included in the cutoff point.

```
cutoff <- function(range) {

  above <- table(fullDist[range,"B_D_A_RipDist"]) +
    table(fullDist[range,"B_D_A_ParDist"]) +
    table(fullDist[range,"B_DDist"]) +
    table(fullDist[range,"D_A_ParDist"])


  below <- table(fullDist[length(range):nrow(ncAvg),"B_D_A_RipDist"]) +
    table(fullDist[length(range):nrow(ncAvg),"B_D_A_ParDist"]) +
    table(fullDist[length(range):nrow(ncAvg),"B_DDist"]) +
    table(fullDist[length(range):nrow(ncAvg),"D_A_ParDist"])

  return((above / (above + below)))
}
```

The annotation function anno() is used as a quick way to annotate the columns of the count data returned from Tophat (they are normally in Enterez format).

```
anno <- function(x) {
  require(AnnotationDbi)
  require(org.Mm.eg.db)

  ### Annotate
  x$symbol <- mapIds(org.Mm.eg.db,
                     keys = row.names(x),
                     column = "SYMBOL",
                     keytype = "ENSEMBL",
                     multiVals = "first")

  ### Remove rows with no gene mapping
  x <- subset(x, select = c(17,1:16))
  x <- x[complete.cases(x), ]
  rownames(x) <- make.names(x$symbol, unique = TRUE)
  x[ ,1] <- NULL

  return(x)
}
```

Load data

The count data (filtered for low counts) is loaded here. Replicates are averaged at the saline

list is ordered by highest counts.

```
nc <- read.csv("finalCountData.csv", header = TRUE)
rownames(nc) <- nc[ ,1]
nc <- nc[ ,-1]

ncAvg <- data.frame(SAL_WT = rowMeans(nc[ ,1:3]),
                    SAL_KO = rowMeans(nc[ ,4:6]),
                    RO_WT = rowMeans(nc[ ,7:9]),
                    RO_KO = nc[ ,10],
                    RO_RAP_WT = rowMeans(nc[ ,11:13]),
                    RO_RAP_KO = rowMeans(nc[ ,14:16]))

ncAvg <- ncAvg[order(ncAvg$SAL_WT, decreasing = TRUE), ]
```

Determining initial cutoff by comparison to other data

We will first be loading all targets from previous data into R. The Miyashiro and Ascano

data will be excluded from the final plotting because it does not have much in common

with our list (similar to what was found by Suhl et al. (Suhl et al., 2014)). However, we are still providing their data in the code below. The consensus overlap lists from *Suhl et al.* are also included below.

```
### Brown data
brownDat <- read.xlsx("brown_list.xls", sheetIndex = 1, header = TRUE,
startRow = 2)
brownProbes <- brownDat[ ,"Probe.Set"] %>% as.vector()
probes <- as.list(mgu74aALIAS2PROBE)
brownGenes <- probes[probes %in% brownProbes] %>% unlist() %>% names()

### Darnell data
darnellDat <- read.xlsx("darnell_list.xls", sheetIndex = 1, header = TR
UE, startRow = 2)
darnellGenes <- darnellDat[ ,"Symbol.from.mm9"] %>% as.vector()

### Ascano PAR-CLIP data
ascanoParDat <- read.xlsx("ascano_par.xlsx", sheetIndex = 1, header = T
RUE, startRow = 3, colIndex = 1:4)
ascanoParGenes <- ascanoParDat[ascanoParDat$Total.mRNA.binding.sites >
0, "Gene"] %>% as.vector()

### Ascano RIP-CHIP data
ascanoRipDat <- read.csv("ascano_rip.csv", header = TRUE, na.strings =
"")
ascanoRipDat <- ascanoRipDat[ ,c(1,3)]
ascanoRipDat <- split(ascanoRipDat, ascanoRipDat$Target.bin)
ascanoRipDat <- ascanoRipDat$Target
ascanoRipGenes <- ascanoRipDat[ ,1] %>% as.vector()

### Miyashiro data
miyashiroDat <- read.csv("miyashiro.csv", header = TRUE)
miyashiroDat <- miyashiroDat[-c(1,1153:1159), ]
miyashiroDat <- miyashiroDat[ ,"UniGene"] %>% as.data.frame()
names(miyashiroDat) <- "unigene"
miyashiroDat$unigene <- as.character(miyashiroDat$unigene)
miyashiroDat$symbol <- mapIds(org.Hs.eg.db,
                              keys = miyashiroDat$unigene,
                              column = "SYMBOL",
                              keytype = "UNIGENE",
                              multiVals = "first")
miyashiroGenes <- miyashiroDat$symbol %>% na.omit() %>% as.vector()
```

```
### Consensus data
BDARIP <- read.xlsx("consensus_FMRP.xlsx", sheetIndex = 2, header = FAL
SE, startRow = 2, stringsAsFactors = FALSE)$X1
BDAPAR <- read.xlsx("consensus_FMRP.xlsx", sheetIndex = 3, header = FAL
SE, startRow = 2, stringsAsFactors = FALSE)$X1
BD <- read.xlsx("consensus_FMRP.xlsx", sheetIndex = 4, header = FALSE,
startRow = 2, stringsAsFactors = FALSE)$X1
DAPAR <- read.xlsx("consensus_FMRP.xlsx", sheetIndex = 5, header = FALS
E, startRow = 2, stringsAsFactors = FALSE)$X1
```

Using the above data, we can now construct a large table that includes all the genes in long format for plotting with ggplot2. The code below also has a few extra details to get the factor levels in the correct order.

```
### Distributions across ranked saline genes for full data sets
fullDist <- data.frame(row.names = row.names(ncAvg) %>% toupper(),
                       order = 1:nrow(ncAvg),
                       brownDist = ifelse((row.names(ncAvg) %>% toupper
()) %in% (brownGenes %>% toupper()), "Brown", NA),
                       darnellDist = ifelse((row.names(ncAvg) %>% toupp
er()) %in% (darnellGenes %>% toupper()), "Darnell", NA),
                       ascanoParDist = ifelse((row.names(ncAvg) %>% tou
pper()) %in% (ascanoParGenes %>% toupper()), "Ascano PAR-CLIP", NA),
                       ascanoRipDist = ifelse((row.names(ncAvg) %>% tou
pper()) %in% (ascanoRipGenes %>% toupper()), "Ascano RIP-CHIP", NA),
                       miyashiroRipDist = ifelse((row.names(ncAvg) %>%
toupper()) %in% (miyashiroGenes %>% toupper()), "Miyashiro", NA),
                       B_D_A_RipDist = ifelse((row.names(ncAvg) %>% tou
pper()) %in% (BDARIP %>% toupper()), "B_D_A-RIP", NA),
                       B_D_A_ParDist = ifelse((row.names(ncAvg) %>% tou
pper()) %in% (BDAPAR %>% toupper()), "B_D_A-PAR", NA),
                       B_DDist = ifelse((row.names(ncAvg) %>% toupper()
) %in% (BD %>% toupper()), "B_D", NA),
                       D_A_ParDist = ifelse((row.names(ncAvg) %>% toupp
er()) %in% (DAPAR %>% toupper()), "D_A-PAR", NA))

fullDist.m <- melt(fullDist, id.vars = "order")

fullDist.m$value <- factor(fullDist.m$value, levels = rev(c("Darnell",
"Brown", "B_D_A-RIP", "B_D_A-PAR", "B_D", "D_A-PAR", "Ascano PAR-CLIP",
"Ascano RIP-CHIP", "Miyashiro")))
fullDist.m$variable <- factor(fullDist.m$variable, levels = rev(c("darn
ellDist", "brownDist", "B_D_A_RipDist", "B_D_A_ParDist", "B_DDist", "D_
```

```
A_ParDist", "ascanoParDist", "ascanoRipDist", "miyashiroRipDist")))

fullDist.m <- fullDist.m[complete.cases(fullDist.m), ]
```

Next, we constructed a cutoff line that was able to capture 95% of the genes within the consensus data list. To do this, we use the cutoff() function described above. For each gene in the data that is included it reports the overlap percentage with the consesus list once the for loop is run. Note that the for loop could be better optimized as it takes quite a while to run. The final data frame orderedCoverage contains all the info as a percentage.

```
### Above/below cutoff line
coverage = NULL
for(i in 1:nrow(ncAvg)) {
  coverage[i] <- cutoff(1:i)
}

orderedCoverage <- data.frame(order = 1:nrow(ncAvg),
                              coverage = coverage * 100)
```

Finally, the data is filtered by including 95% of the genes that overlap with the other data sets. There are two graphs show below. One shows each gene found in our data with the other data sets and shows the cutoff line (shown in red). The bottom graph then shows the percentage of genes that overlap within each row of our ordered data set using the information from orderedCoverage(). The filtered list (4120 genes in total) will now be used moving forward for the next filtration steps.

```
consensusCutoff <- orderedCoverage[orderedCoverage$coverage < 95, ] %>%
nrow()

### Consensus percentage graph
a <- ggplot(orderedCoverage, aes(x = order, y = coverage)) +
  geom_line() +
  geom_vline(xintercept = consensusCutoff, color = "red", linetype = "d
ashed") +
  xlab("Ranked Saline Genes") +
```

```
  ylab("Percent Consensus Lists in Cutoff") +
  theme_bw()

### Consensus overlap graph
b <- ggplot(fullDist.m, aes(x = order, y = value)) +
  geom_point(alpha = 0.75, color = "dodgerblue") +
  geom_vline(xintercept = consensusCutoff, color = "red", linetype = "d
ashed") +
  xlab("") +
  ylab("") +
  theme_bw()

### Fully arranged graphs
gA <- ggplotGrob(a)
gB <- ggplotGrob(b)
grid::grid.newpage()
grid::grid.draw(rbind(gB, gA))
```

Determining KO cutoff by comparison to other data

We have taken two approaches to removing high background counts. The first approach is show here. We know potential targets and non-targets from previous data (Darnell et al., 2011). By determining the enrichment ratio between WT and KO for a given treatment, we can see where distributions of enrichment ratios for the targets or non-targets fall for all the consensus data sets. Using this, we can calibrate and approximate acceptable range of background for our cutoff.

```
### Create enrichment ratios of WT over KO for each treatment
ncAvgDif <- data.frame(row.names = row.names(ncAvg),
                       SAL = ncAvg$SAL_WT / ncAvg$SAL_KO,
                       RO = ncAvg$RO_WT / ncAvg$RO_KO,
                       RO_RAP = ncAvg$RO_RAP_WT / ncAvg$RO_RAP_KO)

### Load targets and non-targets
targets <- c("Gabbr1", "Gabbr2", "Nisch", "Prrc2a", "Kif1a", "Grin2a",
"Adgrb2", "Cyfip2", "Map1b", "Bsn", "Adcy1", "Pde2a", "Lingo1", "Grin2b
")
nontargets <- c("Pabpc1", "Tmem65", "Hprt1", "St8sia3", "Sae1", "Glrb",
```

```
"Gria2", "Tcerg1", "Eif3a", "Slc35f1", "Atp6ap2", "Vldlr")

### Create a data frame that includes the overlap of genes with our dat
a set for each consensus gene list as well as the full Brown and Darnel
l lists
foldRange <- rbind(data.frame(group = "Brown", value = ncAvgDif[row.nam
es(ncAvgDif) %in% brownGenes, ]$SAL),
                   data.frame(group = "Darnell", value = ncAvgDif[row.n
ames(ncAvgDif) %in% darnellGenes, ]$SAL),
                   data.frame(group = "B_D_A-RIP", value = ncAvgDif[(ro
w.names(ncAvgDif) %>% toupper()) %in% BDARIP, ]$SAL),
                   data.frame(group = "B_D_A-PAR", value = ncAvgDif[(ro
w.names(ncAvgDif) %>% toupper()) %in% BDAPAR, ]$SAL),
                   data.frame(group = "B_D", value = ncAvgDif[(row.name
s(ncAvgDif) %>% toupper()) %in% BD, ]$SAL),
                   data.frame(group = "D_A-PAR", value = ncAvgDif[(row.
names(ncAvgDif) %>% toupper()) %in% DAPAR, ]$SAL),
                   data.frame(group = "Non-Targets", value = ncAvgDif[r
ow.names(ncAvgDif) %in% nontargets, ]$SAL))

### Determine average enrichment ratio range between WT and KO for all
groups
foldRange.s <- split(foldRange, foldRange$group)

### Take the mean of the lower quartile of the ranges for all groups (i
ndicated by the dotted line)
cutoff <- sapply(foldRange.s, function(x) quantile(x$value, prob = 0.25
))[1:6] %>% mean()
```

This gives us a cutoff of 1.112788. In other words, a fold enrichment of this value is believed to be the acceptable threshold for where background counts occur. Finally, the data is graphed. We also performed t-tests on all data for the significance bars.

```
### T-tests to show each cutoff is different from non-targets
data.frame(row.names = c("Brown", "Darnell", "B_D_A-RIP", "B_D_A-PAR",
"B_D", "D_A-PAR"),
           p.value = c(t.test(foldRange[foldRange$group == "Brown", "va
lue"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value,
                       t.test(foldRange[foldRange$group == "Darnell", "
value"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value,
                       t.test(foldRange[foldRange$group == "B_D_A-RIP",
"value"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value
,
```

```
                      t.test(foldRange[foldRange$group == "B_D_A-PAR",
"value"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value
,
                      t.test(foldRange[foldRange$group == "B_D", "valu
e"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value,
                      t.test(foldRange[foldRange$group == "D_A-PAR", "
value"], foldRange[foldRange$group == "Non-Targets", "value"])$p.value)
)

##                p.value
## Brown     6.945032e-06
## Darnell   7.145567e-04
## B_D_A-RIP 2.446077e-04
## B_D_A-PAR 7.041381e-06
## B_D       7.660948e-06
## D_A-PAR   4.259868e-04

### Data for signficance bars locations
sig1 <- data.frame(a = c(1, 1:7, 7), b = c(3.4, 3.5, 3.5, 3.5, 3.5, 3.5
, 3.5, 3.5, 3.4))
sig2 <- data.frame(a = c(2, 2:7, 7), b = c(3.2, 3.3, 3.3, 3.3, 3.3, 3.3
, 3.3, 3.2))
sig3 <- data.frame(a = c(3, 3:7, 7), b = c(3.0, 3.1, 3.1, 3.1, 3.1, 3.1
, 3.0))
sig4 <- data.frame(a = c(4, 4:7, 7), b = c(2.8, 2.9, 2.9, 2.9, 2.9, 2.8
))
sig5 <- data.frame(a = c(5, 5:7, 7), b = c(2.6, 2.7, 2.7, 2.7, 2.6))
sig6 <- data.frame(a = c(6, 6:7, 7), b = c(2.4, 2.5, 2.5, 2.4))

### Graph data
c <- ggplot(foldRange, aes(x = group, y = value, fill = group)) +
  geom_boxplot() +
  geom_hline(yintercept = cutoff, color = "red", linetype = "dashed") +
  xlab("") +
  #ylab("WT-to-KO Fold Ratio") +
  ylab("") +
  scale_y_continuous(breaks = c(seq(0,5, by = 0.5))) +
  theme_bw() +
  theme(legend.position = "none")

### Draw signifiance bar lines
c + geom_line(data = sig1, aes(x = a, y = b, fill = NULL)) + annotate("
text", x = 4.0, y = 3.55, label = "***", size = 4) +
  geom_line(data = sig2, aes(x = a, y = b, fill = NULL)) + annotate("te
xt", x = 4.5, y = 3.35, label = "***", size = 4) +
  geom_line(data = sig3, aes(x = a, y = b, fill = NULL)) + annotate("te
xt", x = 5.0, y = 3.15, label = "***", size = 4) +
```

```
  geom_line(data = sig4, aes(x = a, y = b, fill = NULL)) + annotate("te
xt", x = 5.5, y = 2.95, label = "***", size = 4) +
  geom_line(data = sig5, aes(x = a, y = b, fill = NULL)) + annotate("te
xt", x = 6.0, y = 2.75, label = "***", size = 4) +
  geom_line(data = sig6, aes(x = a, y = b, fill = NULL)) + annotate("te
xt", x = 6.5, y = 2.55, label = "***", size = 4)
```

Determining KO cutoff by univariate K-means

We sought to find a more unbiased, quantitative means of assigning a threshold to the background counts. To do this, we used univariate K-means to attempt to separate out a pattern of background WT/KO enrichment from signal from the filtered data set (4120 genes).

```
ncAvgCut <- ncAvg[1:consensusCutoff, ]
ncAvgCut$"Saline" <- ncAvg[1:consensusCutoff, "SAL_WT"] / ncAvg[1:conse
nsusCutoff, "SAL_KO"]
ncAvgCut$"Ro" <- ncAvg[1:consensusCutoff, "RO_WT"] / ncAvg[1:consensusC
utoff, "RO_KO"]
ncAvgCut$"Ro+Rapa" <- ncAvg[1:consensusCutoff, "RO_RAP_WT"] / ncAvg[1:c
onsensusCutoff, "RO_RAP_KO"]

### Univariate k-means clustering. Arrange output in data frame
clustDat <- data.frame(row.names = row.names(ncAvgCut),
                       enriched = ncAvgCut$Saline,
                       cluster = Ckmeans.1d.dp(ncAvgCut$Saline, 2)$clus
ter)

### Define a second cutoff by taking the max in the "background" cluste
r
cutoff2 <- split(clustDat, clustDat$cluster)[[1]]$enriched %>% max()
cutoff2

## [1] 1.289769
```

Visualize WT/KO enrichment ratio with cutoff established

We are now in a position to examine the final KO cutoff using our two separate methods. Here we simply take the average of the two cutoffs for a final cutoff assigned in the variable

KOcutoff. This cutoff will be applied to the data by applying it to the saline fraction only.

The resulting gene list will be our filtered targets.

```
### Define knockout cutoff
KOcutoff <- mean(cutoff, cutoff2)
KOcutoff

## [1] 1.112788

### Filter data to obtain this cutoff
filteredTargets <- ncAvgCut[ncAvgCut$Saline > KOcutoff, 1:6]
```

We now look at the distributions of points that fall above and below the cutoff line.

```
### Create a filter variable that specifies signal or background
ncAvgCut$filter <- ifelse(ncAvgCut$Saline > KOcutoff, "Signal", "Backgr
ound")

### Graph with ggplot2 and arrange together
  ggplot(ncAvgCut, aes(x = SAL_KO %>% log2(), y = SAL_WT %>% log2(), co
lor = filter)) +
    geom_point() +
    scale_color_manual(values = c("grey60", "blue2")) +
    xlab("log2(Knockout)") +
    ylab("log2(Saline)") +
    theme_bw() +
    theme(legend.position = "none")
```

We can also look at the distributions of genes that were removed by the filter compared to the total distribution:

```
ncLog <- ncAvgCut[ ,7:9]
ncLog <- apply(ncLog, 2, log2) %>% as.data.frame()
names(ncLog) <- c("Saline", "Ro", "RoRapa")
ncLog$filter <- ncAvgCut$filter

g1 <- ncLog[ncLog$filter == "Signal", ]
g2 <- ncLog[ncLog$filter == "Background", ]

  ggplot(ncLog, aes(x = Saline)) +
  geom_histogram(data = ncLog[which(ncLog$filter %in% "Signal"), ], aes
```

```
(y = ..density..), fill = "grey60", bins = 100) +
  geom_histogram(data = ncLog[which(ncLog$filter %in% "Background"), ],
aes(y = ..density..), fill = "grey60", bins = 100) +
  geom_density(data = ncLog[which(ncLog$filter %in% "Signal"), ], aes(y
= ..density..), fill = "chartreuse3", alpha = 0.5) +
  geom_density(data = ncLog[which(ncLog$filter %in% "Background"), ], a
es(y = ..density..), fill = "firebrick2", alpha = 0.5) +
  theme_bw()
```

Filtering mitochondrial and glial cells

To remove mitchondrial and glial cells, we searched the MitoMiner (Smith et al., 2012) database and the Brain RNA-seq database for cell types (Zhang et al., 2014)

```
mito <- read.csv("mito.csv", header = FALSE, stringsAsFactors = FALSE)$
V1
glia <- read.csv("glial.csv", header = FALSE, stringsAsFactors = FALSE)
$V1

filteredTargets <- filteredTargets[!(row.names(filteredTargets) %in% c(
mito, glia) ), ]

finalCountData <- nc[row.names(nc) %in% row.names(filteredTargets), ]
```

Final filtration with limma

Finally, we take our filtered list and apply **limma+voom** for differential expression analysis.

```
### Preprocess data columns
countData <- read.csv(file = "combinedCounts.csv", header = TRUE, strin
gsAsFactors = FALSE)
rownames(countData) <- countData[ ,1]
countData[ ,1] <- NULL
cdMain <- subset(countData, select = c(7:12, 16:18, 3:6, 13:15)) %>% an
no()
cdMain <- cdMain[ ,1:9]
cdMain <- cdMain[row.names(cdMain) %in% row.names(finalCountData), ]

### Analyze with limma+voom
```

```
dge <- DGEList(counts = cdMain)
dge <- calcNormFactors(dge)
dge <- cpm(dge, log = TRUE, prior.count = 3)
design <- model.matrix(~ 0+factor(c(1,1,1,2,2,2,3,3,3)))
colnames(design) <- c("RO_RAP_WT", "RO_WT", "SAL_WT")

contrast.matrix <- makeContrasts(SAL_WT-RO_WT, RO_RAP_WT-RO_WT, levels
= design)

v <- voom(cdMain, design, plot = FALSE)
fit <- lmFit(v, design)
fit2 <- contrasts.fit(fit, contrast.matrix)
fit2 <- eBayes(fit2)
resultsRo <- topTable(fit2, coef = 1, number = 25000)
resultsRoRap <- topTable(fit2, coef = 2, number = 25000)

resultsRo$Threshold <- ifelse(resultsRo$adj.P.Val < 0.1, "Significant",
"Non-Significant")
resultsRoRap$Threshold <- ifelse(resultsRo$adj.P.Val < 0.1, "Significan
t", "Non-Significant")
```

Visualization of final dataset

Finally, we can view the cutoffs established by limma with an MA plot, volcano plot, and clustered heatmap.

MA plot

```
ma1 <-
  ggplot(resultsRo, aes(x = AveExpr, y = logFC, color = Threshold)) +
  geom_point(alpha = 0.4, size = 0.8) +
  geom_hline(aes(yintercept = 0), color = "blue", size = 1.2) +
  scale_color_manual(values = c("grey40", "red")) +
  ylim(-1.5,1.5) +
  xlim(6.5,14) +
  xlab("Mean of Normalized Counts") +
  ylab("Log2 Fold-Change") +
  ggtitle("Saline/Ro") +
  theme_bw() +
  theme(legend.position = "none")

ma2 <-
  ggplot(resultsRoRap, aes(x = AveExpr, y = logFC, color = Threshold))
+
  geom_point(alpha = 0.4, size = 0.8) +
```

```
  geom_hline(aes(yintercept = 0), color = "blue", size = 1.2) +
  scale_color_manual(values = c("grey40", "red")) +
  ylim(-1.5,1.5) +
  xlim(6.5,14) +
  xlab("Mean of Normalized Counts") +
  ylab("Log2 Fold-Change") +
  ggtitle("Ro+Rapa/Ro") +
  theme_bw() +
  theme(legend.position = "none")

grid.arrange(ma1, ma2,
             ncol = 2,
             nrow = 1)
```

Volcano plot

```
vp1 <-
  ggplot(resultsRo, aes(x = logFC, y = -log10(adj.P.Val), color = Thres
hold)) +
  geom_point(alpha = 0.4, size = 1.75) +
  xlim(-1,1) +
  ylim(0,2) +
  xlab("Log2 Fold-Change") +
  ylab("-log10 FDR Cutoff") +
  ggtitle("Saline/Ro") +
  theme_bw() +
  theme(legend.position = "none")

vp2 <-
  ggplot(resultsRoRap, aes(x = logFC, y = -log10(adj.P.Val), color = Th
reshold)) +
  geom_point(alpha = 0.4, size = 1.75) +
  xlim(-1.3,1.3) +
  ylim(0,2) +
  xlab("Log2 Fold-Change") +
  ylab("-log10 FDR Cutoff") +
  ggtitle("Ro+Rap/Ro") +
  theme_bw() +
  theme(legend.position = "none")

grid.arrange(vp1, vp2,
             ncol = 2,
             nrow = 1)
```

Clustered heatmap

```
### Heatmap for fold-change data
heatmap <- merge(resultsRo, resultsRoRap, by = 0)
heatmap <- heatmap[heatmap$adj.P.Val.x < 0.1 & heatmap$adj.P.Val.y < 0.
1, ]
heatmap <- heatmap[(heatmap$logFC.x > -2 & heatmap$logFC.x < 2) & (heat
map$logFC.y > -2 & heatmap$logFC.y < 2), ]

heatmap <- data.frame(row.names = heatmap$Row.names,
                      Ro = heatmap$logFC.x,
                      RoRap = heatmap$logFC.y)

pheatmap(heatmap,
         color = brewer.pal(11, "RdYlBu"), fontsize = 7,
         clustering_method = "average",
         clustering_distance_rows = "euclidean",
         clustering_distance_cols = "euclidean",
         show_rownames = FALSE)



### Heatmap for raw count data
finalCountData.log <- apply(finalCountData, 2, log2)
finalCountData.log <- finalCountData.log[row.names(finalCountData.log)
%in% row.names(heatmap), c("SAL_WT1", "SAL_WT2", "SAL_WT3", "RO_WT1", "
RO_WT2", "RO_WT3", "RO_RAP_WT1", "RO_RAP_WT2", "RO_RAP_WT3")]

pheatmap(finalCountData.log,
         color = brewer.pal(11, "RdYlBu"), fontsize = 7,
         clustering_method = "average",
         clustering_distance_rows = "euclidean",
         clustering_distance_cols = "euclidean",
         show_rownames = FALSE)
```

Session info

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8          LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8           LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8       LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8          LC_NAME=en_US.UTF-8
```

```
##  [9] LC_ADDRESS=en_US.UTF-8        LC_TELEPHONE=en_US.UTF-8
## [11] LC_MEASUREMENT=en_US.UTF-8    LC_IDENTIFICATION=en_US.UTF-8
##
## attached base packages:
##  [1] grid      parallel  stats4    stats     graphics  grDevices uti
ls
##  [8] datasets  methods   base
##
## other attached packages:
##  [1] RColorBrewer_1.1-2    pheatmap_1.0.8        Ckmeans.1d.dp_3.4.6
-4
##  [4] edgeR_3.16.4          limma_3.30.5          gridExtra_2.2.1
##  [7] seqinr_3.3-3          org.Hs.eg.db_3.4.0    mgu74a.db_3.2.3
## [10] org.Mm.eg.db_3.4.0    AnnotationDbi_1.36.0  IRanges_2.8.1
## [13] S4Vectors_0.12.0      Biobase_2.34.0        BiocGenerics_0.20.0
## [16] xlsx_0.5.7            xlsxjars_0.6.1        rJava_0.9-8
## [19] magrittr_1.5          biomaRt_2.30.0        reshape2_1.4.2
## [22] ggplot2_2.2.0
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.8      plyr_1.8.4       bitops_1.0-6     tools_3.3.2
##  [5] digest_0.6.10    lattice_0.20-34  evaluate_0.10    RSQLite_1.1
##  [9] memoise_1.0.0    tibble_1.2       gtable_0.2.0     DBI_0.5-1
## [13] yaml_2.1.14      stringr_1.1.0    knitr_1.15.1     locfit_1.5-9
.1
## [17] ade4_1.7-4       rprojroot_1.1    XML_3.98-1.5     rmarkdown_1.
2
## [21] backports_1.0.4  scales_0.4.1     htmltools_0.3.5  assertthat_0
.1
## [25] colorspace_1.3-1 labeling_0.3     stringi_1.1.2    RCurl_1.95-4
.8
## [29] lazyeval_0.2.0   munsell_0.4.3
```

MOTIF ANALYSIS

This section contains all code used to generate the motifs analysis figures.

Load packages

```
library(magrittr)
library(seqinr)
library(lsr)
library(ggplot2)
library(reshape2)
library(biomaRt)
library(Biostrings)
```

Load functions

The getMaxLengthSequences() function is used to find the sequence of maximum length from the list returned from BioMart.

```
getMaxLengthSequences <- function(x) {
  salSeqMax <- split(x, factor(x$mgi_symbol))
  salSeqMax <- lapply(salSeqMax, function(x) max(x$length)) %>% unlist(
) %>% as.data.frame()
  salSeqMax$mgi_symbol <- row.names(salSeqMax)
  row.names(salSeqMax) <- NULL
  names(salSeqMax) <- c("length", "mgi_symbol")
  salSeqUnique <- merge(salSeqMax, x)
  salSeqUnique <- salSeqUnique[order(salSeqUnique$mgi_symbol), ]

  return(salSeqUnique)
}
```

The getmotifs() sequence calculates the occurance of various motifs believed to be present within FMRP target mRNA sequences. For the majority of motifs, they are simply searched using the matchPDict() function from the seqinr package. However, the qfm motif needs its own special function because of its usual properties.

The approach taken here is to first find the two sequences that make up the QFM motif (WGGA = AGGA or AGGT). All instances where they occur are then converted into a dataframe. Next, the occurances of the WGGA motif (start and stop positions) are arranged into a vector and the difference is taken between these start and stop values in order to find the number of nucleotides that are between each occurance of a WGGA. Finally, all occurances less than 6 long (i.e. at least 5 bases between each WGGA) are then returned.

```
getmotifs <- function(x) {
  require(seqinr)

  seq <- x %>% tolower() %>% s2c()
```

```
  qfm <- matchPDict(PDict(c("AGGA", "TGGA")), DNAString(x)) %>% unlist(
) %>% data.frame()
  qfm <- qfm[order(qfm$start), 1:2] %>% as.matrix() %>% t() %>% c() %>%
diff() %>% unique()
  qfm <- sum(qfm[-1] < 6)

  output <- c("ACUK" = count(seq, 4)[c("actg", "actt")] %>% sum(),
              "WGGA" = count(seq, 4)[c("agga", "tgga")] %>% sum(),
              "GAC" = count(seq, 3)["gac"] %>% sum(),
              "GACR" = count(seq, 4)[c("gaca", "gacg")] %>% sum(),
              "GACARG" = count(seq,6)[c("gacaag", "gacagg")] %>% sum(),
              "QFM" = qfm)

  return(output)
}
```

The rho() function in the seqinr package is unable to compute probabilities for multiple motifs at the same time. This function gets around this by manually computing rho for multiple values. This is needed for nearly every motif.

```
multipleRho <- function(sequ, motif1, motif2) {
  require(seqinr)

  freq1 <- count(sequ, nchar(motif1), freq = TRUE)[motif1]
  freq2 <- count(sequ, nchar(motif2), freq = TRUE)[motif2]
  expected1 <- sapply(s2c(motif1), function(x) count(sequ, 1, freq = TR
UE)[unique(x)]) %>% unique() %>% prod()
  expected2 <- sapply(s2c(motif2), function(x) count(sequ, 1, freq = TR
UE)[unique(x)]) %>% unique() %>% prod()

  output <- (freq1 + freq2) / (expected1 + expected2)
  names(output) <- NULL
  return(output)
}
```

The getRho() function calculates rho for all motifs. This function depends on the multipleRho() function described above.

```
getRho <- function(x) {
  require(seqinr)
```

```
  seq <- x %>% tolower() %>% s2c()

  output <- c(ACUK = multipleRho(seq, "actg", "actt"),
              WGGA = multipleRho(seq, "agga", "tgga"),
              GAC = rho(seq, 3)["gac"],
              GACR = multipleRho(seq, "gaca", "gacg"),
              GACARG = multipleRho(seq, "gacaag", "gacagg"))

  names(output) <- c("ACUK", "WGGA", "GAC", "GACR", "GACARG")
  return(output)
}
```

Obtain coding sequences from BioMart

Our analysis consists only of the coding region of FMRP. We have chosen this due to the fact the majority of FMRP binding motifs appear to be clustered in this part of the gene (Anderson et al., 2016). First, we load the data files and obtain the coding sequences:

```
### Load gene list
genes <- read.csv("finalData.csv", header = TRUE, stringsAsFactors = FA
LSE)$Row.names

filteredCounts <- read.csv("finalCountData.csv", header = TRUE, strings
AsFactors = FALSE, row.names = 1) %>% row.names()

### Get all sequences for the filtered list
ensembl <- useMart("ensembl", dataset = "mmusculus_gene_ensembl")

sequ <- biomaRt::getSequence(id = filteredCounts,
                             type = "mgi_symbol",
                             seqType = "coding",
                             mart = ensembl)
```

Next, we filter the sequence list into two categories: either "target" which consists of the final filtered limma list and "non-target" which represent those that are not in that list.

```
### Remove unavailable sequences
sequ <- sequ[!grepl("Sequence unavailable", sequ$'coding', fixed = TRUE
), ]

### Find the length of each sequence then pick the max length sequence
```

```
if there are multiple sequences returned from BioMart
sequ <- sequ[order(sequ$mgi_symbol), ]
sequ$length <- sapply(sequ$coding, nchar)
sequ <- getMaxLengthSequences(sequ)

### Filter into either "targets" or "non-targets"
targets <- sequ[sequ$mgi_symbol %in% genes, ]
nonTargets <- sequ[!(sequ$mgi_symbol %in% genes), ]
```

Analyze coding sequences for motifs

Warning! Running the first line of code is a very computationally intensive process. A matrix is generated that calculates all motifs for every sequence string in the data set.

```
### Targets
targetMotifs <- matrix(sapply(targets[ ,"coding"], getmotifs), nrow = n
row(targets), ncol = 6, byrow = TRUE)
targetMotifTable <- data.frame(Length = targets$length)
rownames(targetMotifTable) <- make.names(targets$mgi_symbol, unique = T
RUE)
targetMotifTable <- cbind(targetMotifTable, targetMotifs) %>% as.data.f
rame()
names(targetMotifTable) <- c("Length", "ACUK", "WGGA", "GAC", "GACR", "
GACARG", "QFM")

targetMotifTable$ACUK.freq <- round((targetMotifTable$ACUK * 1000) / ta
rgetMotifTable$Length, 2)
targetMotifTable$WGGA.freq <- round((targetMotifTable$WGGA * 1000) / ta
rgetMotifTable$Length, 2)
targetMotifTable$GAC.freq <- round((targetMotifTable$GAC * 1000) / targ
etMotifTable$Length, 2)
targetMotifTable$GACR.freq <- round((targetMotifTable$GACR * 1000) / ta
rgetMotifTable$Length, 2)
targetMotifTable$GACARG.freq <- round((targetMotifTable$GACARG * 1000)
/ targetMotifTable$Length, 2)
targetMotifTable$QFM.freq <- round((targetMotifTable$QFM * 1000) / targ
etMotifTable$Length, 2)

rhoMotifs <- matrix(sapply(targets[, "coding"], getRho), nrow = nrow(ta
rgets), ncol = 5, byrow = TRUE)

### Non-targets
nonTargetMotifs <- matrix(sapply(nonTargets[ ,"coding"], getmotifs), nr
ow = nrow(nonTargets), ncol = 6, byrow = TRUE)
```

```
nonTargetMotifTable <- data.frame(Length = nonTargets$length)
row.names(nonTargetMotifTable) <- make.names(nonTargets$mgi_symbol, uni
que = TRUE)
nonTargetMotifTable <- cbind(nonTargetMotifTable, nonTargetMotifs) %>%
as.data.frame()
names(nonTargetMotifTable) <- c("Length", "ACUK", "WGGA", "GAC", "GACR"
, "GACARG", "QFM")

nonTargetMotifTable$ACUK.freq <- round((nonTargetMotifTable$ACUK * 1000
) / nonTargetMotifTable$Length, 2)
nonTargetMotifTable$WGGA.freq <- round((nonTargetMotifTable$WGGA * 1000
) / nonTargetMotifTable$Length, 2)
nonTargetMotifTable$GAC.freq <- round((nonTargetMotifTable$GAC * 1000)
/ nonTargetMotifTable$Length, 2)
nonTargetMotifTable$GACR.freq <- round((nonTargetMotifTable$GACR * 1000
) / nonTargetMotifTable$Length, 2)
nonTargetMotifTable$GACARG.freq <- round((nonTargetMotifTable$GACARG *
1000) / nonTargetMotifTable$Length, 2)
nonTargetMotifTable$QFM.freq <- round((nonTargetMotifTable$QFM * 1000)
/ nonTargetMotifTable$Length, 2)

rhoMotifsNonTarget <- matrix(sapply(nonTargets[, "coding"], getRho), nr
ow = nrow(nonTargets), ncol = 5, byrow = TRUE)
```

Calculate statistics

```
### T-tests
tt <- c(t.test(targetMotifTable$ACUK.freq, nonTargetMotifTable$ACUK.fre
q)$p.value,
        t.test(targetMotifTable$WGGA.freq, nonTargetMotifTable$WGGA.fre
q)$p.value,
        t.test(targetMotifTable$GAC.freq, nonTargetMotifTable$GAC.freq)
$p.value,
        t.test(targetMotifTable$GACR.freq, nonTargetMotifTable$GACR.fre
q)$p.value,
        t.test(targetMotifTable$GACARG.freq, nonTargetMotifTable$GACARG
.freq)$p.value,
        t.test(targetMotifTable$QFM, nonTargetMotifTable$QFM)$p.value)

### Wilcox-Mann_Whitley tests
wmw <- c(wilcox.test(targetMotifTable$ACUK.freq, nonTargetMotifTable$AC
UK.freq, correct = FALSE)$p.value,
         wilcox.test(targetMotifTable$WGGA.freq, nonTargetMotifTable$WG
GA.freq, correct = FALSE)$p.value,
         wilcox.test(targetMotifTable$GAC.freq, nonTargetMotifTable$GAC
.freq, correct = FALSE)$p.value,
         wilcox.test(targetMotifTable$GACR.freq, nonTargetMotifTable$GA
```

```
CR.freq, correct = FALSE)$p.value,
        wilcox.test(targetMotifTable$GACARG.freq, nonTargetMotifTable$
GACARG.freq, correct = FALSE)$p.value,
        wilcox.test(targetMotifTable$QFM, nonTargetMotifTable$QFM, cor
rect = FALSE)$p.value)

### Effect size
cd <- c(cohensD(targetMotifTable$ACUK.freq, nonTargetMotifTable$ACUK.fr
eq),
        cohensD(targetMotifTable$WGGA.freq, nonTargetMotifTable$WGGA.fr
eq),
        cohensD(targetMotifTable$GAC.freq, nonTargetMotifTable$GAC.freq
),
        cohensD(targetMotifTable$GACR.freq, nonTargetMotifTable$GACR.fr
eq),
        cohensD(targetMotifTable$GACARG.freq, nonTargetMotifTable$GACAR
G.freq),
        cohensD(targetMotifTable$QFM, nonTargetMotifTable$QFM))
```

Assemble final motif table

```
motifSummary <- data.frame("Mean Frequency per kB (Cutoff)" = colMeans(
targetMotifTable)[8:13],
                           "Mean Frequency per kB (Non-Cutoff)" = colMe
ans(nonTargetMotifTable)[8:13],
                           "Rho (Cutoff)" = c(colMeans(rhoMotifs), NA),
                           "Rho (Non-Cutoff)" = c(colMeans(rhoMotifsNon
Target), NA),
                           "t-test" = round(tt, 3),
                           "Wilcoxon-Mann-Whitney Test" = round(wmw, 3)
,
                           "Cohen's d" = round(cd, 3))

motifSummary
```

```
##             Mean.Frequency.per.kB..Cutoff.
## ACUK.freq                        8.0357233
## WGGA.freq                       16.5308086
## GAC.freq                        16.7872366
## GACR.freq                        7.9048368
## GACARG.freq                      0.8750705
## QFM.freq                         0.6477374
##             Mean.Frequency.per.kB..Non.Cutoff. Rho..Cutoff.
## ACUK.freq                             8.4833728    0.4547952
## WGGA.freq                            16.6661166    0.1999265
## GAC.freq                             16.0662057    0.9553620
## GACR.freq                             7.4436598    0.2222723
```

```
## GACARG.freq                          0.7731510    0.0243765
## QFM.freq                             0.6544778           NA
##              Rho..Non.Cutoff. t.test Wilcoxon.Mann.Whitney.Test Cohen
.s.d
## ACUK.freq         0.47667503  0.000                      0.000     0
.144
## WGGA.freq         0.20053122  0.222                      0.239     0
.026
## GAC.freq          0.93583759  0.000                      0.000     0
.162
## GACR.freq         0.21408773  0.000                      0.000     0
.152
## GACARG.freq       0.02209649  0.000                      0.000     0
.115
## QFM.freq                  NA  0.000                      0.000     0
.088
```

Visualize motifs

Visualization was performed similar to (Suhl et al., 2014). This is simply the visualization of the table produced above.

Motif frequency

```
motifFreq.m <- rbind(melt(targetMotifTable[ ,8:12]),
                     melt(nonTargetMotifTable[ ,8:12])) %>% as.data.frame()
motifFreq.m$type <- c(rep("Target", nrow(melt(targetMotifTable[ ,8:12])
)),
                      rep("Non-Target", nrow(melt(nonTargetMotifTable[ ,8:1
2]))))
levels(motifFreq.m$variable) <- c("ACUK", "WGGA", "GAC", "GACR", "GACAR
G")
motifFreq.m$type <- factor(motifFreq.m$type, levels = c("Target", "Non-
Target"))

  ggplot(motifFreq.m, aes(x = type, y = value, fill = type)) +
  geom_boxplot() +
  facet_grid(. ~ variable) +
  xlab("") +
  ylab("") +
  labs(fill = "") +
  scale_fill_manual(values = c("dodgerblue4", "darkgoldenrod2")) +
  theme_bw()
```

QFM analysis

```
qfmCount.m <- rbind(melt(targetMotifTable[ ,7]),
                    melt(nonTargetMotifTable[ ,7])) %>% as.data.frame(
)
qfmCount.m$type <- c(rep("Target", nrow(melt(targetMotifTable[ ,7]))),
                     rep("Non-Target", nrow(melt(nonTargetMotifTable[
,7]))))
qfmCount.m$type <- factor(qfmCount.m$type, levels = c("Target", "Non-Ta
rget"))

  ggplot(qfmCount.m, aes(x = type, y = value, fill = type)) +
  geom_boxplot() +
  scale_fill_manual(values = c("dodgerblue4", "darkgoldenrod2")) +
  xlab("") +
  ylab("") +
  theme_bw()
```

Session info

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats4    parallel  stats     graphics  grDevices utils     data
sets
## [8] methods   base
##
## other attached packages:
##  [1] Biostrings_2.42.0   XVector_0.14.0      IRanges_2.8.1
##  [4] S4Vectors_0.12.0    BiocGenerics_0.20.0 biomaRt_2.30.0
##  [7] reshape2_1.4.2      ggplot2_2.2.0       lsr_0.5
## [10] seqinr_3.3-3        magrittr_1.5
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.8          plyr_1.8.4           bitops_1.0-6
##  [4] tools_3.3.2          zlibbioc_1.20.0      digest_0.6.10
```

```
##  [7] evaluate_0.10        RSQLite_1.1          memoise_1.0.0
## [10] tibble_1.2           gtable_0.2.0         DBI_0.5-1
## [13] yaml_2.1.14          stringr_1.1.0        knitr_1.15.1
## [16] rprojroot_1.1        ade4_1.7-4           grid_3.3.2
## [19] Biobase_2.34.0       AnnotationDbi_1.36.0 XML_3.98-1.5
## [22] rmarkdown_1.2        backports_1.0.4      scales_0.4.1
## [25] htmltools_0.3.5      assertthat_0.1       colorspace_1.3-1
## [28] labeling_0.3         stringi_1.1.2        RCurl_1.95-4.8
## [31] lazyeval_0.2.0       munsell_0.4.3
```

CLUSTERING ANALYSIS

This section contains all the data used for clustering the filtered count data.

Load packages

```
library(ggplot2)
library(reshape2)
library(magrittr)
library(RDAVIDWebService)
library(org.Mm.eg.db)
library(gridExtra)
library(xlsx)
library(factoextra)
library(Ckmeans.1d.dp)
library(dplyr)
library(gdata)
library(amap)
```

Load functions

The functionalAnnotationChart() function a list of genes, and ontology (term), and an output file name. It returns the Functional Annotation Chart from DAVID.

```
functionalAnnotationChart <- function(genes, term, filename) {

  require(RDAVIDWebService)
  require(org.Mm.eg.db)

  david <- DAVIDWebService$new(email = "snamjoshi87@utexas.edu", url =
"https://david.ncifcrf.gov/webservice/services/DAVIDWebService.DAVIDWeb
ServiceHttpSoap12Endpoint/")
  egenes <- select(org.Mm.eg.db, genes, "ENTREZID", "SYMBOL")[,2]
```

```
  chart <- addList(david, egenes,
                   idType = "ENTREZ_GENE_ID",
                   listName = "Functional Annotation Chart",
                   listType = "Gene")

  setAnnotationCategories(david, term)
  getFunctionalAnnotationChartFile(david, fileName = filename)

  return(paste("Ouput DAVID analysis file:", filename))
}
```

The uniqueFilteredGoTerms() function a list of downregulated genes and a list of upregulated genes from DAVID results, and a desired PValue and FDR cutoff. It first applies these cutoffs to both data sets. Then, it finds the unique GO terms between the remaining terms.

```
uniqueFilteredGoTerms <- function(downfileName, upfileName, PValue, FDR
) {
  downBP <- read.table(downfileName, sep = "\t", header = TRUE, strings
AsFactors = FALSE, quote = "")
  downBP <- downBP[ ,c("Term", "PValue", "FDR")]
  downBP <- downBP[downBP$PValue < PValue | downBP$FDR < FDR, ]

  upBP <- read.table(upfileName, sep = "\t", header = TRUE, stringsAsFa
ctors = FALSE, quote = "")
  upBP <- upBP[ , c("Term", "PValue", "FDR")]
  upBP <- upBP[upBP$PValue < PValue | upBP$FDR < FDR, ]

  sharedBP <- intersect(downBP$Term, upBP$Term)
  downOnlyBP <- setdiff(downBP$Term, upBP$Term)
  upOnlyBP <- setdiff(upBP$Term, downBP$Term)

  downOnlyTableBP <- downBP[downBP$Term %in% downOnlyBP, c("Term", "PVa
lue", "FDR")]
  upOnlyTableBP <- upBP[upBP$Term %in% upOnlyBP, c("Term", "PValue", "F
DR")]

  downOnlyTableBP <- downOnlyTableBP[order(downOnlyTableBP$PValue), "Te
rm"]
  upOnlyTableBP <- upOnlyTableBP[order(upOnlyTableBP$PValue), "Term"]

  return(list(downTable = downOnlyTableBP, upTable = upOnlyTableBP))
}
```

The simplifyGo() function takes a DAVID functional annotation chart and creates a similarity matrix for each GO term with every other GO term by using two different similarity metrics: the jaccard cutoff (JC) and overlap coefficient (OC) where $JC = \frac{|A \cap B|}{|A \cup B|}$ and $OC = \frac{|A \cap B|}{Min(|A|,|B|)}$. This approach is based off of (Merico et al., 2010; Merico et al., 2011). The similarity metrics are calculated within the subfunction called similarity(). These metrics are two different ways of assessing the degree of overlap between two gene lists.

The average score of JC and OC for each set of genes associated with a GO term is computed in a giant similarity matrix for all GO terms. Note that this procedure is very computational intensive. Finally, the matrix is filtered for a similarity cutoff of > 0.5. The function then returns this matrix.

```
simplifyGo <- function(david_chart) {

  ### Similarity function for overlap coefficient and jaccard cutoff
  similarity <- function(list) {

    A <- list[[1]]
    B <- list[[2]]

    JC <- function(A, B) {
      output <- (intersect(A, B) %>% length()) / (union(A, B) %>% lengt
h())
      return(output)
    }

    OC <- function(A, B) {
      output <- (intersect(A, B) %>% length()) / (min(length(A), length
(B)))
      return(output)
    }

    similarity <- mean(c(OC(A, B), JC(A, B)))
    return(similarity)
```

```
  }

  ### Create similarity matrix
  list1 <- david_chart[ ,"Genes"] %>% strsplit(split  = ",") %>% lapply
(FUN = as.numeric)
  list2 <- list1

  computeSimilarity <- Vectorize(function(x, y) {
    vec1 <- list1[[x]]
    vec2 <- list2[[y]]

    similarity(list(vec1, vec2))
  })

  filterMatrix <- outer(1:length(list1), 1:length(list2), computeSimila
rity)
  rownames(filterMatrix) <- david_chart[ ,"Term"]
  colnames(filterMatrix) <- david_chart[ ,"Term"]
  filterMatrix[filterMatrix < 0.5] <- 0

  return(filterMatrix)
}
```

The getDendrogramClusters() function is used to extract the cluster data from the hcut() function.

```
getDendrogramClusters <- function(hcut_output) {
  clusters <- data.frame(Cluster = hcut_output$cluster %>% unname(),
                         Term = hcut_output$labels)

  labels <- merge(data.frame(Term = labels(hcut_output %>% as.dendrogra
m()),
                             Order = 1:nrow(clusters)), clusters, "Term
")

  output <- labels[order(labels$Order), ]
  return(output)
}
```

The determineClusterNumbers() function utilizes the within sum of squares (WSS) to estimate cluster number for hierachical clustering of the GO terms (based on their similarity index). The function returns a plot which is used to approximate the clusters.

```
determineClusterNumbers <- function(x, n) {
  require(ggplot2)

  wss <- (nrow(x) - 1) * sum(apply(x, 2, var))

  for(i in 2:n) {
    wss[i] <- sum(kmeans(x, centers = i)$withinss)
  }

  wss <- as.data.frame(cbind(wss, clusters = 1:n))

  plot <- ggplot(wss, aes(clusters, wss)) +
    geom_line() +
    geom_point() +
    theme_bw()

  return(plot)
}
```

The determineClusterNumbers1D() function operates simarilary to the above function but instead provides WSS for a univariate distribution with K-means.

```
determineClusterNumbers1D <- function(x, n) {
  require(ggplot2)
  require(Ckmeans.1d.dp)

  wss <- (length(x) - 1) * var(x)

  for(i in 2:n) {
    wss[i] <- sum(Ckmeans.1d.dp(x, i)$withinss)
  }

  wss <- as.data.frame(cbind(wss, clusters = 1:n))

  plot <- ggplot(wss, aes(clusters, wss)) +
    geom_line() +
    geom_point() +
    theme_bw()

  return(plot)
}
```

The kMeanPlot() function takes a two column data frame (x) and the number of cluster (n), clusters the data, and graphs it.

```
require(amap)

colorFunc <- colorRampPalette(c("lightblue", "blue"))

kMeansPlot <- function(x, n) {
  clustDat <- as.matrix(x)
  clustDat[is.na(clustDat)] <- 0
  kClust <- Kmeans(clustDat, centers = n, method = "pearson", iter.max
= 50)

  clustDat <- as.data.frame(cbind(clustDat, clusters = kClust$cluster))
  clustDat$clusters <- as.factor(clustDat$clusters)
  centers <- as.data.frame(kClust$centers)

  clustPlot <- ggplot(clustDat, aes(RO, RO_RAP, color = clusters, fill
= clusters)) +
    geom_point() +
    geom_vline(xintercept = 0, color = "black") +
    geom_hline(yintercept = 0, color = "black") +
    xlab("Ro Fold Change") +
    ylab("Ro+Rapamycin Fold Change") +
    xlim(-1.5,1.5) +
    ylim(-1.5,1.5) +
    ggtitle("K Means Clustering for Ro or Ro+Rapa Fold Changes") +
    theme_bw()

  clustPlot
}
```

Load data and filter

```
fc <- read.csv("finalData.csv", header = TRUE, stringsAsFactors = FALSE
)
row.names(fc) <- fc$Row.names
fc[ ,1] <- NULL
fc <- fc[ ,c("logFC.x", "logFC.y")]
names(fc) <- c("RO", "RO_RAP")

counts <- read.csv("finalCountData.csv", header = TRUE, stringsAsFactor
s = FALSE, row.names = 1)
countsAvg <- data.frame(row.names = row.names(counts),
                        SAL_WT = rowMeans(counts[ ,1:3]),
```

```
                        RO_WT = rowMeans(counts[ ,7:9]),
                        RO_RAP_WT = rowMeans(counts[ ,11:13]))
countsAvg.log <- apply(countsAvg, 2, log2) %>% as.data.frame()

### Here we apply a fold-change filter
roDownGenes <- fc[fc$RO < quantile(fc$RO, 0.25), ]
roUpGenes <- fc[fc$RO > quantile(fc$RO, 0.75), ]

roRapDownGenes <- fc[fc$RO_RAP < quantile(fc$RO_RAP, 0.25), ]
roRapUpGenes <- fc[fc$RO_RAP > quantile(fc$RO_RAP, 0.75), ]
```

Differential binding

This graph is just to illustrate the differential binding effects seen in different fold-change conditions.

```
p1 <-
  ggplot(countsAvg.log, aes(x = SAL_WT, y = RO_WT)) +
  geom_point(color = "dodgerblue3") +
  xlab("Normalized Saline Counts") +
  ylab("Normalized Ro Counts") +
  theme_bw()

p2 <-
  ggplot(countsAvg.log, aes(x = SAL_WT, y = RO_RAP_WT)) +
  geom_point(color = "dodgerblue3") +
  xlab("Normalized Saline Counts") +
  ylab("Normalized Ro + Rapamycin Counts") +
  theme_bw()

p3 <-
  ggplot(countsAvg.log, aes(x = RO_RAP_WT, y = RO_WT)) +
  geom_point(color = "dodgerblue3") +
  xlab("Normalized Ro + Rapamycin Counts") +
  ylab("Normalized Ro Counts") +
  theme_bw()

grid.arrange(p1, p2, p3,
             ncol = 3,
             nrow = 1)
```

Unique GO clustering by fold-change group

Code to generate DAVID Functional Annotation Charts for Ro and RoRap. For all DAVID charts, we used a p-value cutoff of 0.001 and FDR cutoff of 0.05. The strategy here is to use the GO cutoffs from early to cluster on the basis of similarity. This will group like GO-terms together in a branch on the dendrogram. This unbiased approach allows us to give our own names manually to each major cluster family. Then, we can take the average p-value within each branch, the average number of genes per GO term, and plot this with **ggplot2**. To determine cluster number, the within sum of squares (WSS) is graphed for *n* clusters. The chosen cluster number attempts to minimize the WSS of squares by visually seeing where the overall shape of the plot begins to flatten out on the y-axis.

```
### Query DAVID for Ro and get unique GO terms
functionalAnnotationChart(roDownGenes,
                          term = "GOTERM_BP_ALL",
                          filename = "david_ro_down_1000_biological_pro
cess.txt")

functionalAnnotationChart(roUpGenes,
                          term = "GOTERM_BP_ALL",
                          filename = "david_ro_up_1000_biological_proce
ss.txt")

uniqueBpRo <- uniqueFilteredGoTerms(downfileName = "david_ro_down_1000_
biological_process.txt",
                                    upfileName = "david_ro_up_1000_biol
ogical_process.txt",
                                    PValue = 0.001,
                                    FDR = 0.05)

### Query DAVID for Ro+Rap and get unique GO terms
functionalAnnotationChart(roRapDownGenes,
                          term = "GOTERM_BP_ALL",
                          filename = "david_roRap_down_1000_biological_
process.txt")
```

```
functionalAnnotationChart(roRapUpGenes,
                          term = "GOTERM_BP_ALL",
                          filename = "david_roRap_up_1000_biological_pr
ocess.txt")

uniqueBpRoRap <- uniqueFilteredGoTerms(downfileName = "david_roRap_down
_1000_biological_process.txt",
                                       upfileName = "david_roRap_up_100
0_biological_process.txt",
                                       PValue = 0.001,
                                       FDR = 0.05)

### Query DAVID for candidate depression genes
dep <- read.csv("depression_gene_candidates.csv", header = FALSE, strin
gsAsFactors = FALSE)$V1

dep <- sapply(dep, function(x) {
  y <- tolower(x)
  y <- paste0(toupper(substr(y, 1, 1)), substr(y, 2, nchar(y)))
  return(y)
  })

names(dep) <- NULL

functionalAnnotationChart(dep,
                          term = "GOTERM_BP_ALL",
                          filename = "david_depression_1000_biological_
process.txt")

### Query DAVID for consensus FMRP genes
BD <- read.xlsx("consensus_FMRP.xlsx", sheetIndex = 4, header = FALSE,
startRow = 2, stringsAsFactors = FALSE)$X1

BD <- sapply(BD, function(x) {
  y <- tolower(x)
  y <- paste0(toupper(substr(y, 1, 1)), substr(y, 2, nchar(y)))
  return(y)
})

names(BD) <- NULL

functionalAnnotationChart(BD,
                          term = "GOTERM_BP_ALL",
                          filename = "david_BD_1000_biological_process.
txt")
```

Ro top GO chart

```
### Load DAVID data and filter
roUp <- read.table("david_ro_up_1000_biological_process.txt", header =
TRUE, sep = "\t", stringsAsFactors = FALSE)
roUp <- roUp[roUp$PValue < 0.001 | roUp$FDR < 0.05, c("Term", "PValue",
"Genes", "Count")]
roDown <- read.table("david_ro_down_1000_biological_process.txt", heade
r = TRUE, sep = "\t", stringsAsFactors = FALSE, quote = "")
roDown <- roDown[roDown$PValue < 0.001 | roDown$FDR < 0.05, c("Term", "
PValue", "Genes", "Count")]

roUp <- roUp[!(roUp$Term %in% roDown$Term), ]
roDown <- roDown[!(roDown$Term %in% roUp$Term), ]

### Determine optimal cluster number
roUpMatrix <- simplifyGo(roUp)
determineClusterNumbers(roUpMatrix, 30)



### Create dendrogram and get clusters
roUpDendrogram <- hcut(roUpMatrix, k = 30, stand = TRUE)
fviz_dend(roUpDendrogram, rect = TRUE, cex = 0.5, show_labels = TRUE)



roUpClusters <- getDendrogramClusters(roUpDendrogram)

### Merge clusters with original DAVID data
roUpClusters.mg <- merge(roUpClusters, roUp, "Term")[ ,c("Term", "PValu
e", "Count", "Cluster", "Order")]
roUpClusters.mg$NegLogPValue <- -log10(roUpClusters.mg$PValue)
roUpClusters.mg$Direction <- rep("Up", nrow(roUpClusters.mg))
roUpClusters.mg <- roUpClusters.mg[order(roUpClusters.mg$Order), ]

# write.csv to export file
```

Now, the dendrogram data is manually annotated and re-imported to R.

```
roUpChart <- read.csv("roUpClusters_network_data.csv", header = TRUE, s
tringsAsFactors = FALSE)

### Create Ro Chart
roUpChart.s <- split(roUpChart, roUpChart$Category)
roUpSummary <- data.frame(Term = names(roUpChart.s),
                          NegLogPValue = lapply(roUpChart.s, function(x
```

```
) mean(x$NegLogPValue)) %>% unlist() %>% unname(),
                           Count = lapply(roUpChart.s, function(x) mean(
x$Count)) %>% unlist() %>% unname(),
                           stringsAsFactors = FALSE)
roUpSummary <- roUpSummary[roUpSummary$NegLogPValue > -log(0.01), ]
roUpSummary$Term <- factor(roUpSummary$Term, levels = roUpSummary[order
(roUpSummary$NegLogPValue, decreasing = FALSE), "Term"])

ggplot(roUpSummary, aes(x = NegLogPValue, y = Term, size = Count)) +
  geom_segment(aes(yend = Term), xend = 0, color = "black", size = 0.5)
+
  geom_point(shape = 21, color = "black", fill = "dodgerblue3") +
  scale_size_area(max_size = 5.5) +
  scale_size_continuous(limits = c(1, 160), breaks = c(40, 80, 120, 160
)) +
  scale_x_continuous(limits = c(4, 16), breaks = seq(4, 16, by = 2)) +
  ylab("") +
  xlab("-log(p-value)") +
  theme_bw() +
  theme(axis.text.x = element_text(color = "grey30", size = 12),
        axis.text.y = element_text(color = "grey30", size = 12),
        axis.title.x = element_text(size = 17))
```

Ro+Rap top GO chart

```
### Load DAVID data and filter
roRapUp <- read.table("david_roRap_up_1000_biological_process.txt", hea
der = TRUE, sep = "\t", stringsAsFactors = FALSE)
roRapUp <- roRapUp[roRapUp$PValue < 0.001 | roRapUp$FDR < 0.05, c("Term
", "PValue", "Genes", "Count")]
roRapDown <- read.table("david_roRap_down_1000_biological_process.txt",
header = TRUE, sep = "\t", stringsAsFactors = FALSE)
roRapDown <- roRapDown[roRapDown$PValue < 0.001 | roRapDown$FDR < 0.05,
c("Term", "PValue", "Genes", "Count")]

roRapUp <- roRapUp[!(roRapUp$Term %in% roRapDown$Term), ]

### Determine optimal cluster number
roRapUpMatrix <- simplifyGo(roRapUp)
determineClusterNumbers(roRapUpMatrix, 30)



### Create dendrogram and get clusters
roRapUpDendrogram <- hcut(roRapUpMatrix, k = 30, stand = TRUE)
```

```
fviz_dend(roRapUpDendrogram, rect = TRUE, cex = 0.5, show_labels = TRUE
) # 803 x 469



roRapUpClusters <- getDendrogramClusters(roRapUpDendrogram)

### Merge clusters with original DAVID data
roRapUpClusters.mg <- merge(roRapUpClusters, roRapUp, "Term")[ ,c("Term
", "PValue", "Count", "Cluster", "Order")]
roRapUpClusters.mg$NegLogPValue <- -log10(roRapUpClusters.mg$PValue)
roRapUpClusters.mg$Direction <- rep("Up", nrow(roRapUpClusters.mg))
roRapUpClusters.mg <- roRapUpClusters.mg[order(roRapUpClusters.mg$Order
), ]

# write.csv to export file
```

Now, the dendrogram data is manually annotated and re-imported to R.

```
### Create RoRap Chart
roRapUpChart <- read.csv("roRapUpClusters_network_data.csv", header = T
RUE, stringsAsFactors = FALSE)
roRapUpChart.s <- split(roRapUpChart, roRapUpChart$Category)
roRapUpSummary <- data.frame(Term = names(roRapUpChart.s),
                             NegLogPValue = lapply(roRapUpChart.s, functio
n(x) mean(x$NegLogPValue)) %>% unlist() %>% unname(),
                             Count = lapply(roRapUpChart.s, function(x) me
an(x$Count)) %>% unlist() %>% unname(),
                             stringsAsFactors = FALSE)
roRapUpSummary <- roRapUpSummary[roRapUpSummary$NegLogPValue > -log(0.0
1), ]
roRapUpSummary$Term <- factor(roRapUpSummary$Term, levels = roRapUpSumm
ary[order(roRapUpSummary$NegLogPValue, decreasing = FALSE), "Term"])

ggplot(roRapUpSummary, aes(x = NegLogPValue, y = Term, size = Count)) +
  geom_segment(aes(yend = Term), xend = 0, color = "black", size = 0.5)
+
  geom_point(shape = 21, color = "black", fill = "dodgerblue3") +
  scale_size_area(max_size = 5.5) +
  scale_size_continuous(limits = c(1, 300), breaks = c(50, 100, 150, 20
0, 250, 300)) +
  scale_x_continuous(limits = c(4, 24), breaks = seq(4, 24, by = 4)) +
  ylab("") +
  xlab("-log(p-value)") +
  theme_bw() +
  theme(axis.text.x = element_text(color = "grey30", size = 12),
```

```
        axis.text.y = element_text(color = "grey30", size = 12),
        axis.title.x = element_text(size = 17))
```

Depression top GO chart

```
depChart <- read.csv("depression_network_data.csv", header = TRUE, stri
ngsAsFactors = FALSE)
depChart.s <- split(depChart, depChart$Category)
depChartSummary <- data.frame(Term = names(depChart.s),
                              NegLogPValue = lapply(depChart.s, funct
ion(x) mean(x$NegLogPValue)) %>% unlist() %>% unname(),
                              Count = lapply(depChart.s, function(x)
mean(x$Count)) %>% unlist() %>% unname(),
                              stringsAsFactors = FALSE)

depChartSummary$Term <- factor(depChartSummary$Term, levels = depChartS
ummary[order(depChartSummary$NegLogPValue, decreasing = FALSE), "Term"]
)

ggplot(depChartSummary[-34, ], aes(x = NegLogPValue, y = Term, size = C
ount)) +
  geom_segment(aes(yend = Term), xend = 0, color = "black", size = 0.5)
+
  geom_point(shape = 21, color = "black", fill = "grey50") +
  scale_size_area(max_size = 4) +
  ggtitle("Enriched GO Biological Process Terms") +
  ylab("") +
  xlab("-log(p-value") +
  theme_bw()
```

FMRP consensus top GO chart

```
bdChart <- read.csv("BD_network_data.csv", header = TRUE, stringsAsFact
ors = FALSE)
bdChart.s <- split(bdChart, bdChart$Category)
bdChartSummary <- data.frame(Term = names(bdChart.s),
                             NegLogPValue = lapply(bdChart.s, function
(x) mean(x$NegLogPValue)) %>% unlist() %>% unname(),
                             Count = lapply(bdChart.s, function(x) mea
n(x$Count)) %>% unlist() %>% unname(),
                             stringsAsFactors = FALSE)

bdChartSummary$Term <- factor(bdChartSummary$Term, levels = bdChartSumm
```

```
ary[order(bdChartSummary$NegLogPValue, decreasing = FALSE), "Term"])

ggplot(bdChartSummary, aes(x = NegLogPValue, y = Term, size = Count)) +
  geom_segment(aes(yend = Term), xend = 0, color = "black", size = 0.5)
+
  geom_point(shape = 21, color = "black", fill = "#FF66FF") +
  scale_size_area(max_size = 4) +
  ggtitle("Enriched GO Biological Process Terms") +
  ylab("") +
  xlab("-log(p-value") +
  theme_bw()
```

Univariate K-means clustering and GO ontology

Ro fold-changes

```
determineClusterNumbers1D(fc$RO, 15)

roClust <- Ckmeans.1d.dp(fc$RO, 8)
roClusters <- data.frame(row.names = row.names(fc),
                         foldChange = fc$RO,
                         cluster = roClust$cluster)   ### Gather cluste
ring data

### Remove tails of distribution
roClusters <- roClusters[roClusters$foldChange < 1.5, ]
roClusters <- roClusters[roClusters$foldChange > -1.6, ]
```

For GO clustering, we took the genes from each cluster and sent them to DAVID. For example, in the first cluster we would use the following code:

```
roClusters.s <- split(roClusters, roClusters$cluster)
functionalAnnotationChart(roClusters.s[[1]] %>% row.names(),
                          term = "GOTERM_BP_ALL",
                          filename = "DAVID_BP_KM1D_CLUST1.txt")
```

The GO clustering reveals that for 1-D clustering across the entire distribution by fold-change, the fifth cluster contains genes enriched for trans-synaptic signaling that change together (average p-value = $6.00 * 10^{-8}$).

```
### For loop to set boundaries of clusters within distribution during p
lotting
roBoundaries = NULL

for(i in 1:(roClusters$cluster %>% max())) {
  roBoundaries <- rbind(roBoundaries, split(roClusters, roClusters$clus
ter)[[i]]$foldChange %>% range())
}

### Plot data
p <-
  ggplot(roClusters, aes(x = foldChange)) +
    geom_density() +
    theme_bw()

d <- ggplot_build(p)$data[[1]]

p +
  geom_area(data = subset(d, x < roBoundaries[7,2]), aes(x = x, y = y),
fill = "grey92") +
  geom_area(data = subset(d, x < roBoundaries[6,2]), aes(x = x, y = y),
fill = "grey92") +
  geom_area(data = subset(d, x < roBoundaries[5,2]), aes(x = x, y = y),
fill = "red") +
  geom_area(data = subset(d, x < roBoundaries[4,2]), aes(x = x, y = y),
fill = "grey92") +
  geom_area(data = subset(d, x < roBoundaries[3,2]), aes(x = x, y = y),
fill = "grey92") +
  geom_area(data = subset(d, x < roBoundaries[2,2]), aes(x = x, y = y),
fill = "grey92") +
  geom_area(data = subset(d, x < roBoundaries[1,2]), aes(x = x, y = y),
fill = "grey92")
```

Ro+Rap fold-changes

```
determineClusterNumbers1D(fc$RO_RAP, 15)



roRapClust <- Ckmeans.1d.dp(fc$RO_RAP, 8)
roRapClusters <- data.frame(row.names = row.names(fc),
                            foldChange = fc$RO_RAP,
                            cluster = roRapClust$cluster)   ### Gather clu
stering data
```

```
### Remove tails of distribution
roRapClusters <- roRapClusters[roRapClusters$foldChange < 1.5, ]
roRapClusters <- roRapClusters[roRapClusters$foldChange > -1.6, ]

### For loop to set boundaries of clusters within distribution during p
lotting
roRapBoundaries = NULL

for(i in 1:(roRapClusters$cluster %>% max())) {
  roRapBoundaries <- rbind(roRapBoundaries, split(roRapClusters, roRapC
lusters$cluster)[[i]]$foldChange %>% range())
}

### Plot data
p1 <-
  ggplot(roRapClusters, aes(x = foldChange)) +
    geom_density() +
    theme_bw()

d1 <- ggplot_build(p1)$data[[1]]

p1 +
  geom_area(data = subset(d1, x < roRapBoundaries[7,2]), aes(x = x, y =
y), fill = "grey92") +
  geom_area(data = subset(d1, x < roRapBoundaries[6,2]), aes(x = x, y =
y), fill = "grey92") +
  geom_area(data = subset(d1, x < roRapBoundaries[5,2]), aes(x = x, y =
y), fill = "red") +
  geom_area(data = subset(d1, x < roRapBoundaries[4,2]), aes(x = x, y =
y), fill = "grey92") +
  geom_area(data = subset(d1, x < roRapBoundaries[3,2]), aes(x = x, y =
y), fill = "grey92") +
  geom_area(data = subset(d1, x < roRapBoundaries[2,2]), aes(x = x, y =
y), fill = "grey92") +
  geom_area(data = subset(d1, x < roRapBoundaries[1,2]), aes(x = x, y =
y), fill = "grey92")
```

Fold-Change Correlation

The purpose of this plot is to see gene that changed in the same direction between the two calculated fold-change categories (Saline/Ro or Ro+Rap/Ro).

```
### Load data for trans-synaptic genes so they can be uniquely colored
on the plot.
transRo <- read.csv("roTranssynaptic.csv", stringsAsFactors = FALSE)
transRo[ ,1] <- NULL
transRoRap <- read.csv("roRapTranssynaptic.csv", stringsAsFactors = FAL
SE)
transRoRap[ ,1] <- NULL
trans <- c(transRo$genes, transRoRap$genes) %>% unique()

fc$Category <- rep("No Change", nrow(fc))
fc[fc$RO > quantile(fc$RO, 0.75) | fc$RO_RAP > quantile(fc$RO_RAP, 0.75
), "Category"] <- "Up"
fc[fc$RO < quantile(fc$RO, 0.25) | fc$RO_RAP < quantile(fc$RO_RAP, 0.25
), "Category"] <- "Down"
fc[row.names(fc) %in% trans, "Category" ] <- "Transsynaptic"
plotOrder <- c("No Change", "Down", "Up", "Transsynaptic")
fc$Category <- reorder.factor(fc$Category, new.order = plotOrder)
fc <- fc %>% arrange(Category)

ggplot(fc, aes(x = RO_RAP, y = RO, color = Category, fill = Category))
+
  geom_point(shape = 21) +
  geom_hline(yintercept = 0) +
  geom_vline(xintercept = 0) +
  scale_x_continuous(limits = c(-1.5, 1.5)) +
  scale_y_continuous(limits = c(-1.5, 1.5)) +
  xlab("Log2 Fold Change (Ro+Rap / Ro)") +
  ylab("Log2 Fold Change (Saline / Ro)") +
  scale_fill_manual(values = c("grey50", "deepskyblue1", "deepskyblue1"
, "red2")) +
  scale_color_manual(values = c("black", "black", "black", "black")) +
  theme_bw() +
  theme(axis.text.x = element_text(color = "black", size = 12),
        axis.text.y = element_text(color = "black", size = 12),
        axis.title.x = element_text(size = 17),
        axis.title.y = element_text(size = 17))
```

Bivariate K-means clustering and GO ontology

Finally, we will cluster the above genes using 2-D K-means. We will use the same approach as before by minimizing the WSS for clusters.

```
### Load data and cluster
fc <- read.csv("finalData.csv", header = TRUE, stringsAsFactors = FALSE
)
row.names(fc) <- fc$Row.names
fc[ ,1] <- NULL
fc <- fc[ ,c("logFC.x", "logFC.y")]
names(fc) <- c("RO", "RO_RAP")

determineClusterNumbers(fc, 15)



kMeansPlot(fc, 11)



### Extract clusters
clusterID <- data.frame(gene = row.names(fc),
                        cluster = as.integer(kMeansPlot(fc, 8)$data$clu
sters),
                        stringsAsFactors = FALSE)
clusterID <- split(clusterID, clusterID$cluster)

### Perform GO clustering as before
```

Session info

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
##  [1] LC_CTYPE=en_US.UTF-8          LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8           LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8       LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8          LC_NAME=en_US.UTF-8
##  [9] LC_ADDRESS=en_US.UTF-8        LC_TELEPHONE=en_US.UTF-8
## [11] LC_MEASUREMENT=en_US.UTF-8    LC_IDENTIFICATION=en_US.UTF-8
##
## attached base packages:
## [1] stats4    parallel  stats     graphics  grDevices utils     data
sets
## [8] methods   base
##
## other attached packages:
##  [1] amap_0.8-14             gdata_2.17.0
##  [3] dplyr_0.5.0             Ckmeans.1d.dp_3.4.6-4
##  [5] factoextra_1.0.3        xlsx_0.5.7
```

```
##  [7] xlsxjars_0.6.1          rJava_0.9-8
##  [9] gridExtra_2.2.1         org.Mm.eg.db_3.4.0
## [11] RDAVIDWebService_1.12.0 GOstats_2.40.0
## [13] Category_2.40.0         Matrix_1.2-7.1
## [15] AnnotationDbi_1.36.0    IRanges_2.8.1
## [17] S4Vectors_0.12.0        Biobase_2.34.0
## [19] graph_1.52.0            BiocGenerics_0.20.0
## [21] magrittr_1.5            reshape2_1.4.2
## [23] ggplot2_2.2.0
##
## loaded via a namespace (and not attached):
##  [1] mclust_5.2              ggrepel_0.6.5          Rcpp_0.12.8
##  [4] mvtnorm_1.0-5           lattice_0.20-34        GO.db_3.4.0
##  [7] class_7.3-14            gtools_3.5.0           assertthat_0.1
## [10] rprojroot_1.1           digest_0.6.10          R6_2.2.0
## [13] plyr_1.8.4              backports_1.0.4        RSQLite_1.1
## [16] evaluate_0.10           diptest_0.75-7         lazyeval_0.2.0
## [19] annotate_1.52.0         kernlab_0.9-25         whisker_0.3-2
## [22] rmarkdown_1.2           labeling_0.3           splines_3.3.2
## [25] stringr_1.1.0           RCurl_1.95-4.8         munsell_0.4.3
## [28] htmltools_0.3.5         nnet_7.3-12            tibble_1.2
## [31] dendextend_1.3.0        XML_3.98-1.5           AnnotationForge_1
.16.0
## [34] MASS_7.3-45             bitops_1.0-6           grid_3.3.2
## [37] RBGL_1.50.0             xtable_1.8-2           GSEABase_1.36.0
## [40] gtable_0.2.0            DBI_0.5-1              scales_0.4.1
## [43] stringi_1.1.2           genefilter_1.56.0      flexmix_2.3-13
## [46] robustbase_0.92-6       tools_3.3.2            fpc_2.1-10
## [49] trimcluster_0.1-2       DEoptimR_1.0-8         survival_2.40-1
## [52] yaml_2.1.14             colorspace_1.3-1       cluster_2.0.5
## [55] prabclus_2.2-6          memoise_1.0.0          knitr_1.15.1
## [58] modeltools_0.2-21
```

SHINY APP FOR PARK7 BINDING MOTIF IDENTIFICATION

This document contains all the code needed to run the Shiny app that determines Park7 binding sites.

User interface

```
library(shiny)

shinyUI(fluidPage(

  titlePanel("Determine Putative Park7 RNA Targets"),
```

```
  sidebarLayout(

    sidebarPanel(width = 4,
      selectInput(inputId = "dataChoice",
                  label = "1. Select data type:",
                  choices = c("Gene and sequence list" = "geneSeq",
                              "Gene list only" = "gene")),

      fileInput(inputId = "dat",
                label = "2. Load data (.csv):",
                accept = ".csv"),


      conditionalPanel(
        condition = "input.dataChoice == 'gene'",
        radioButtons(inputId = "biomartType",
                     label = "3. Retrieve sequence from Biomart:",
                     choices = c("5utr", "cds", "3utr"),
                     inline = TRUE)

      ),

      actionButton(inputId = "execute",
                   label = "Find binding sites and view data",
                   style = "color: #fff; background-color: #337ab7; bor
der-color: #2e6da4"),


      tags$hr(),


      conditionalPanel(
        condition = "input.dataChoice == 'gene'",

        tags$div(class = "header", checked = NA,
                 tags$b("4. Download data: ")),

        actionButton(inputId = "biomartDat",
                     label = "Biomart sequence data",
                     icon("download")),

        actionButton(inputId = "bindingDat",
                     label = "Binding site data",
                     icon("download"))
```

```
      ),

      conditionalPanel(
        condition = "input.dataChoice == 'geneSeq'",

        tags$div(class = "header", checked = NA,
                 tags$b("3. Download data: ")),

        actionButton(inputId = "bindingDat2",
                     label = "Binding site data",
                     icon("download"))
      ),

      tags$hr(),

      tags$div(class = "header", checked = NA,
               tags$b("Instructions:")),

      tags$div(class = "header", checked = NA,
               tags$p("First select your data type. A gene list should
be a
                      file that contains a single column of mouse gene
s in the
                      official gene symbol format ONLY and no column h
eader.
                      A gene and sequence list contains a gene list in
the first
                      column, a list of sequences in the second column
, and no
                      column header. All input files must be in .csv f
ormat.
                      If your input is a gene list, the sequences
                      will be retrieved automatically using the Ensemb
le Biomart
                      service. The resulting table is displayed on the
right and
                      can be manipulated to sort by a specified column
.")),

      tags$hr(),

      tags$div(class = "header", checked = NA,
               tags$p("This tool was developed by the following members
of the
                      Raab-Graham Lab: Sanjeev Namjoshi, Farr Niere, J
uan Penaranda,
```

```
                        Kimberly Raab-Graham, and Cameron Reynoldson. Co
ncept developed by
                        S.N., F.N., and K.R.G based on Park7 CLIP data [
1]. Original code
                        written in Python by J.P. and C.R. Code translat
ed into R and Shiny
                        app developed by S.N. For any questions or comme
nts please email S.N.
                        at snamjoshi87@utexas.edu. All code is available
on GitHub."))

    ),
    mainPanel(
      DT::dataTableOutput("table")
    )
  )
))
```

Server

```
library(shiny)
library(DT)

shinyServer(function(input, output) {
  rv <- reactiveValues()
  rv$seq <- NULL
  rv$bioMart <- NULL

  observe({
    if(input$execute) {
      indat <- input$dataChoice
      infile <- input$dat

      if(indat == "geneSeq") {
        seq <- read.csv(infile$datapath, stringsAsFactors = FALSE, head
er = FALSE)
        names(seq) <- c("sequence", "gene")

        withProgress(message = "Finding binding sites...",
                     value = NULL, {
                       rv$seq <- createTableOutput(seq)
                       incProgress(0.1)
                     })
      }

      if(indat == "gene") {
        genes <- read.csv(infile$datapath, stringsAsFactors = FALSE, he
```

```
ader = FALSE)
        type <- input$biomartType

        withProgress(message = "Contacting BioMart...",
                     value = NULL, {
                       seqBioMart <- get_sequences(genes, type)
                     })

        names(seqBioMart) <- c("sequence", "gene")
        rv$bioMart <- seqBioMart

        withProgress(message = "Finding binding sites...",
                     value = NULL, {
                       rv$seq <- createTableOutput(seqBioMart)
                     })
      }
    }

    if(input$biomartDat > 0) {
      fname <- paste("Biomart_data_", input$biomartType, Sys.Date(), ".
csv", sep = '')

      withProgress(message = "Downloading BioMart data...",
                   value = NULL, {
                     write.csv(rv$bioMart, fname)
                   })
    }

    if(input$bindingDat > 0) {
      fname <- paste("Binding_data", Sys.Date(), ".csv", sep = '')
      withProgress(message = "Downloading BioMart data...",
                   value = NULL, {
                     write.csv(rv$seq, fname)
                   })

    }

    if(input$bindingDat2 > 0) {
      fname <- paste("Binding_data", Sys.Date(), ".csv", sep = '')
      withProgress(message = "Downloading BioMart data...",
                   value = NULL, {
                     write.csv(rv$seq, fname)
                   })
    }
  })
```

```
  output$table <- DT::renderDataTable(DT::datatable({
    rv$seq
  }))
})
```

Global functions

```
### Load packages
library(biomaRt)
library(magrittr)
library(Biostrings)
library(seqinr)

windows <- c("CNGCNG", "GNGCNG")

get_sequences <- function(genes, type) {

  ensembl <- useMart("ensembl", dataset = "mmusculus_gene_ensembl")

  if(type == "5utr") {
    seq <- getSequence(id = genes,
                       type = "mgi_symbol",
                       seqType = "5utr",
                       mart = ensembl)

    seq <- seq[!grepl("Sequence unavailable", seq$'5utr', fixed = TRUE)
,]
  }

  if(type == "cds") {
    seq <- getSequence(id = genes,
                        type = "mgi_symbol",
                        seqType = "coding",
                        mart = ensembl)
  }

  if(type == "3utr") {
    seq <- getSequence(id = genes,
                       type = "mgi_symbol",
                       seqType = "3utr",
                       mart = ensembl)

    seq <- seq[!grepl("Sequence unavailable", seq$'3utr', fixed = TRUE)
,]
  }

  return(seq)
```

```
}

createTableOutput <- function(seq) {
  ### Calculate Sequence length
  seqlength <- sapply(seq$sequence, nchar) %>% unlist() %>% unname()

  ### Calculate CXGCXG indices and matches
  w1Full <- sapply(seq$sequence, function(x) matchPattern(windows[1], D
NAString(x), fixed = FALSE))
  w1Full <- lapply(w1Full, function(x) slot(x, "ranges") %>% as.data.fr
ame())
  w1Matches <- sapply(w1Full, nrow) %>% unlist() %>% unname()
  w1Indices <- lapply(w1Full, function(x) paste(x$start, collapse = ",
")) %>% unlist()

  ### Calculate GXGCXG indices and matches
  w2Full <- sapply(seq$sequence, function(x) matchPattern(windows[2], D
NAString(x), fixed = FALSE))
  w2Full <- lapply(w2Full, function(x) slot(x, "ranges") %>% as.data.fr
ame())
  w2Matches <- sapply(w2Full, nrow) %>% unlist() %>% unname()
  w2Indices <- lapply(w2Full, function(x) paste(x$start, collapse = ",
")) %>% unlist()

  ### Calculate frequency per kB
  freq <- ((1000 * (w1Matches + w2Matches)) / seqlength) %>% round(2)

  ### Generate table
  output <- data.frame("Gene" = seq$gene,
                       "Seq.Length" = seqlength,
                       "CXGCXG Indices" = w1Indices,
                       "CXGCXG Matches" = w1Matches,
                       "GXGCXG Indices" = w2Indices,
                       "GXGCXG Matches" = w2Matches,
                       "Total Matches" = w1Matches + w2Matches,
                       "Freq/kB" = freq,
                       stringsAsFactors = FALSE)

}
```

Session info

```
## R version 3.3.2 (2016-10-31)
## Platform: x86_64-pc-linux-gnu (64-bit)
## Running under: Ubuntu 16.04.1 LTS
##
## locale:
```

```
##  [1] LC_CTYPE=en_US.UTF-8       LC_NUMERIC=C
##  [3] LC_TIME=en_US.UTF-8        LC_COLLATE=en_US.UTF-8
##  [5] LC_MONETARY=en_US.UTF-8    LC_MESSAGES=en_US.UTF-8
##  [7] LC_PAPER=en_US.UTF-8       LC_NAME=C
##  [9] LC_ADDRESS=C               LC_TELEPHONE=C
## [11] LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## loaded via a namespace (and not attached):
##  [1] backports_1.0.4 magrittr_1.5    rprojroot_1.1   tools_3.3.2
##  [5] htmltools_0.3.5 yaml_2.1.14     Rcpp_0.12.8     stringi_1.1.2
##  [9] rmarkdown_1.2   knitr_1.15.1    stringr_1.1.0   digest_0.6.10
## [13] evaluate_0.10
```

Appendix E: Prior Publications

Namjoshi, S., Heaney, C.F., and Raab-Graham, K.F. [*In preparation*]. mTOR mediates trans-synaptic signaling required for synapse formation and rapid antidepression efficacy requiring Fragile X Mental Retardation Protein. [Chapter 4]

Cacheaux, L.P., Niere, F., **Namjoshi, S.**, Reynoldson, C.J., Shah, Y., Zemelman, B.V., Raab-Graham, K.F. [*In preparation*]. Parkinson protein 7 links aberrant calcium dynamics and glutamate receptor trafficking to models of autism through coordinated expression of α2∂2.

Wolfe, S.A., Farris, S.P., Heaney, C.F., Erikson, E., **Namjoshi, S.**, Mayfield, R.D., Harris, R.A., Raab-Graham, K.F. [*In preparation*]. Rapid antidepressant action of an NMDAR antagonist and alcohol suggests overlapping differential exon usage for synapse remodeling.

Namjoshi, S. and Raab-Graham, K.F. (2017). Screening the Molecular Framework Underlying Local Dendritic mRNA Translation. *Front. Mol. Neurosci*. 10(45). doi: 10.3389/fnmol.2017.00045. [Chapter 1]

Wolfe, S.A., Workman, E.R., Heaney, C., Niere, F., **Namjoshi, S.**, Cacheaux, L.P., Farris, S.P., Drew, M.R., Zemelman, B.V., Harris, R.A., Raab-Graham, K.F. (2016). FMRP Regulates an Ethanol-dependent Shift in $GABA_BR$ function with Rapid Antidepressant Properties. *Nat. Commun. 7: 12867*. doi: 10.1038/ncomms12867. [Appendix B]

Raab-Graham, K.F., Workman, E.R., **Namjoshi, S.**, Niere, F. (2016). Pushing the threshold: How NMDAR antagonists induce homeostasis through protein synthesis to remedy depression. *Brain Res*. 1647: 94-104. doi: 10.1016/j.brainres.2016.04.020. [Appendix A]

Niere, F., **Namjoshi, S.**, Song E., Dilly, G.A., Schoenhard, G., Zemelman, B.V., Mechref, Y., Raab-Graham, K.F. (2016). Analysis of Proteins That Rapidly Change Upon Mechanistic/Mammalian Target of Rapamycin Complex 1 (mTORC1) Repression Identifies Parkinson Protein 7 (PARK7) as a Novel Protein Aberrantly Expressed in Tuberous Sclerosis Complex (TSC). *Mol Cell Proteomics* 15(2): 426-44. doi: 10.1074/mcp.M115.055079. [Chapters 2 and 3]

Biblography

Aakalu, G., Smith, W.B., Nguyen, N., Jiang, C., and Schuman, E.M. (2001). Dynamic visualization of local protein synthesis in hippocampal neurons. *Neuron* 30**,** 489-502.

Abdallah, C.G., Sanacora, G., Duman, R.S., and Krystal, J.H. (2015). Ketamine and rapid-acting antidepressants: a window into a new neurobiology for mood disorder therapeutics. *Annu Rev Med* 66**,** 509-523.

Abelaira, H.M., Reus, G.Z., Neotti, M.V., and Quevedo, J. (2014). The role of mTOR in depression and antidepressant responses. *Life Sci* 101**,** 10-14.

Abraham, M.T., Kuriakose, M.A., Sacks, P.G., Yee, H., Chiriboga, L., Bearer, E.L., and Delacure, M.D. (2001). Motility-related proteins as markers for head and neck squamous cell cancer. *Laryngoscope* 111**,** 1285-1289.

Abrahams, B.S., Arking, D.E., Campbell, D.B., Mefford, H.C., Morrow, E.M., Weiss, L.A., Menashe, I., Wadkins, T., Banerjee-Basu, S., and Packer, A. (2013). SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Mol Autism* 4**,** 36.

Aebersold, R., Burlingame, A.L., and Bradshaw, R.A. (2013). Western blots versus selected reaction monitoring assays: time to turn the tables? *Mol Cell Proteomics* 12**,** 2381-2382.

Afgan, E., Baker, D., Van Den Beek, M., Blankenberg, D., Bouvier, D., Cech, M., Chilton, J., Clements, D., Coraor, N., Eberhard, C., Gruning, B., Guerler, A., Hillman-Jackson, J., Von Kuster, G., Rasche, E., Soranzo, N., Turaga, N., Taylor, J., Nekrutenko, A., and Goecks, J. (2016). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Res* 44**,** W3-W10.

Ainsley, J.A., Drane, L., Jacobs, J., Kittelberger, K.A., and Reijmers, L.G. (2014). Functionally diverse dendritic mRNAs rapidly associate with ribosomes following a novel experience. *Nat Commun* 5**,** 4510.

Akinfiresoye, L., and Tizabi, Y. (2013). Antidepressant effects of AMPA and ketamine combination: role of hippocampal BDNF, synapsin, and mTOR. *Psychopharmacology (Berl)* 230**,** 291-298.

Aksoy-Aksel, A., Zampa, F., and Schratt, G. (2014). MicroRNAs and synaptic plasticity--a mutual relationship. *Philos Trans R Soc Lond B Biol Sci* 369.

Aldinger, K.A., Plummer, J.T., Qiu, S., and Levitt, P. (2011). SnapShot: genetics of autism. *Neuron* 72**,** 418-418 e411.

An, W.L., Cowburn, R.F., Li, L., Braak, H., Alafuzoff, I., Iqbal, K., Iqbal, I.G., Winblad, B., and Pei, J.J. (2003). Up-regulation of phosphorylated/activated p70 S6 kinase and its relationship to neurofibrillary pathology in Alzheimer's disease. *Am J Pathol* 163**,** 591-607.

Anders, S., Mccarthy, D.J., Chen, Y., Okoniewski, M., Smyth, G.K., Huber, W., and Robinson, M.D. (2013). Count-based differential expression analysis of RNA sequencing data using R and Bioconductor. *Nat Protoc* 8**,** 1765-1786.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31**,** 166-169.

Anderson, B.R., Chopra, P., Suhl, J.A., Warren, S.T., and Bassell, G.J. (2016). Identification of consensus binding sites clarifies FMRP binding determinants. *Nucleic Acids Res* 44**,** 6649-6659.

Anderson, M.P. (2010). Arrested glutamatergic synapse development in human partial epilepsy. *Epilepsy Curr* 10**,** 153-158.

Ando, R., Hama, H., Yamamoto-Hino, M., Mizuno, H., and Miyawaki, A. (2002). An optical marker based on the UV-induced green-to-red photoconversion of a fluorescent protein. *Proc Natl Acad Sci U S A* 99**,** 12651-12656.

Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data* [Online]. Available: http://www.bioinformatics.babraham.ac.uk/projects/fastqc [Accessed].

Antar, L.N., Li, C., Zhang, H., Carroll, R.C., and Bassell, G.J. (2006). Local functions for FMRP in axon growth cone motility and activity-dependent regulation of filopodia and spine synapses. *Mol Cell Neurosci* 32**,** 37-48.

Antion, M.D., Hou, L., Wong, H., Hoeffer, C.A., and Klann, E. (2008a). mGluR-dependent long-term depression is associated with increased phosphorylation of S6 and synthesis of elongation factor 1A but remains expressed in S6K-deficient mice. *Mol Cell Biol* 28**,** 2996-3007.

Antion, M.D., Merhav, M., Hoeffer, C.A., Reis, G., Kozma, S.C., Thomas, G., Schuman, E.M., Rosenblum, K., and Klann, E. (2008b). Removal of S6K1 and S6K2 leads to divergent alterations in learning, memory, and synaptic plasticity. *Learn Mem* 15**,** 29-38.

Antonucci, F., Corradini, I., Morini, R., Fossati, G., Menna, E., Pozzi, D., Pacioni, S., Verderio, C., Bacci, A., and Matteoli, M. (2013). Reduced SNAP-25 alters short-term plasticity at developing glutamatergic synapses. *EMBO Rep* 14**,** 645-651.

Aoto, J., Nam, C.I., Poon, M.M., Ting, P., and Chen, L. (2008). Synaptic signaling by all-trans retinoic acid in homeostatic synaptic plasticity. *Neuron* 60**,** 308-320.

Aragam, N., Wang, K.S., and Pan, Y. (2011). Genome-wide association analysis of gender differences in major depressive disorder in the Netherlands NESDA and NTR population-based samples. *J Affect Disord* 133**,** 516-521.

Arancio, O., Kiebler, M., Lee, C.J., Lev-Ram, V., Tsien, R.Y., Kandel, E.R., and Hawkins, R.D. (1996). Nitric oxide acts directly in the presynaptic neuron to produce long-term potentiation in cultured hippocampal neurons. *Cell* 87**,** 1025-1035.

Ariga, H., Takahashi-Niki, K., Kato, I., Maita, H., Niki, T., and Iguchi-Ariga, S.M. (2013). Neuroprotective function of DJ-1 in Parkinson's disease. *Oxid Med Cell Longev* 2013**,** 683920.

Arnold, J.B. (2016). "ggthemes: Extra Themes, Scales and Geoms for 'ggplot2'". R package version 3.3.0 ed.).

Ascano, M., Hafner, M., Cekan, P., Gerstberger, S., and Tuschl, T. (2012a). Identification of RNA-protein interaction networks using PAR-CLIP. *Wiley Interdiscip Rev RNA* 3**,** 159-177.

Ascano, M., Jr., Mukherjee, N., Bandaru, P., Miller, J.B., Nusbaum, J.D., Corcoran, D.L., Langlois, C., Munschauer, M., Dewell, S., Hafner, M., Williams, Z., Ohler, U., and Tuschl, T. (2012b). FMRP targets distinct mRNA sequence elements to regulate protein expression. *Nature* 492**,** 382-386.

Association, A.P. (2013). *Diagnostic and statistical manual of mental disorders* Arlington, Virginia: American Psychiatric Publishing.

Auer, P.L., and Doerge, R.W. (2010). Statistical design and analysis of RNA sequencing data. *Genetics* 185**,** 405-416.

Auerbach, B.D., Osterweil, E.K., and Bear, M.F. (2011). Mutations causing syndromic autism define an axis of synaptic pathophysiology. *Nature* 480**,** 63-68.

Auguie, B. (2016). "gridExtra: Miscellaneous Functions for "Grid" Graphics". R package version 2.2.1 ed.).

Autry, A.E., Adachi, M., Nosyreva, E., Na, E.S., Los, M.F., Cheng, P.F., Kavalali, E.T., and Monteggia, L.M. (2011). NMDA receptor blockade at rest triggers rapid behavioural antidepressant responses. *Nature* 475**,** 91-95.

Bache, S.M., and Wickham, H. (2014). "magrittr: A Forward-Pipe Operator for R". R package version 1.5 ed.).

Bachis, A., Cruz, M.I., Nosheny, R.L., and Mocchetti, I. (2008). Chronic unpredictable stress promotes neuronal apoptosis in the cerebral cortex. *Neurosci Lett* 442**,** 104-108.

Ballou, L.M., and Lin, R.Z. (2008). Rapamycin and mTOR kinase inhibitors. *J Chem Biol* 1**,** 27-36.

Baltz, A.G., Munschauer, M., Schwanhausser, B., Vasile, A., Murakawa, Y., Schueler, M., Youngs, N., Penfold-Brown, D., Drew, K., Milek, M., Wyler, E., Bonneau, R., Selbach, M., Dieterich, C., and Landthaler, M. (2012). The mRNA-bound proteome and its global occupancy profile on protein-coding transcripts. *Mol Cell* 46**,** 674-690.

Banasr, M., and Duman, R.S. (2008). Glial loss in the prefrontal cortex is sufficient to induce depressive-like behaviors. *Biol Psychiatry* 64**,** 863-870.

Banerjee, S., Neveu, P., and Kosik, K.S. (2009). A coordinated local translational control point at the synapse involving relief from silencing and MOV10 degradation. *Neuron* 64**,** 871-884.

Barnett, D.G., and Bibb, J.A. (2011). The role of Cdk5 in cognition and neuropsychiatric and neurological pathology. *Brain Res Bull* 85**,** 9-13.

Baruzzo, G., Hayer, K.E., Kim, E.J., Di Camillo, B., Fitzgerald, G.A., and Grant, G.R. (2017). Simulation-based comprehensive benchmarking of RNA-seq aligners. *Nat Methods* 14**,** 135-139.

Bassell, G.J., and Warren, S.T. (2008). Fragile X syndrome: loss of local mRNA regulation alters synaptic development and function. *Neuron* 60**,** 201-214.

Bateup, H.S., Johnson, C.A., Denefrio, C.L., Saulnier, J.L., Kornacker, K., and Sabatini, B.L. (2013). Excitatory/inhibitory synaptic imbalance leads to hippocampal hyperexcitability in mouse models of tuberous sclerosis. *Neuron* 78**,** 510-522.

Bateup, H.S., Takasaki, K.T., Saulnier, J.L., Denefrio, C.L., and Sabatini, B.L. (2011). Loss of Tsc1 in vivo impairs hippocampal mGluR-LTD and increases excitatory synaptic function. *J Neurosci* 31**,** 8862-8869.

Batista, G., Johnson, J.L., Dominguez, E., Costa-Mattioli, M., and Pena, J.L. (2016). Translational control of auditory imprinting and structural plasticity by eIF2alpha. *Elife* 5.

Bayes, A., Collins, M.O., Croning, M.D., Van De Lagemaat, L.N., Choudhary, J.S., and Grant, S.G. (2012). Comparative study of human and mouse postsynaptic proteomes finds high compositional conservation and abundance differences for key synaptic proteins. *PLoS One* 7**,** e46683.

Bear, M.F. (1995). Mechanism for a sliding synaptic modification threshold. *Neuron* 15**,** 1-4.

Bekinschtein, P., Cammarota, M., Igaz, L.M., Bevilaqua, L.R., Izquierdo, I., and Medina, J.H. (2007). Persistence of long-term memory storage requires a late protein synthesis- and BDNF- dependent phase in the hippocampus. *Neuron* 53**,** 261-277.

Bengtsson, H. (2016). "R.utils: Various Programming Utilities". R package version 2.5.0 ed.).

Bennett, M.R., and Lagopoulos, J. (2014). Stress and trauma: BDNF control of dendritic-spine formation and regression. *Prog Neurobiol* 112**,** 80-99.

Benowitz, L.I., and Routtenberg, A. (1997). GAP-43: an intrinsic determinant of neuronal development and plasticity. *Trends Neurosci* 20**,** 84-91.

Beretta, L., Gingras, A.C., Svitkin, Y.V., Hall, M.N., and Sonenberg, N. (1996). Rapamycin blocks the phosphorylation of 4E-BP1 and inhibits cap-dependent initiation of translation. *EMBO J* 15**,** 658-664.

Berman, R.M., Cappiello, A., Anand, A., Oren, D.A., Heninger, G.R., Charney, D.S., and Krystal, J.H. (2000). Antidepressant effects of ketamine in depressed patients. *Biol Psychiatry* 47**,** 351-354.

Bermejo, M.K., Milenkovic, M., Salahpour, A., and Ramsey, A.J. (2014). Preparation of synaptic plasma membrane and postsynaptic density proteins using a discontinuous sucrose gradient. *J Vis Exp***,** e51896.

Bertram, L., Mcqueen, M.B., Mullin, K., Blacker, D., and Tanzi, R.E. (2007). Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 39**,** 17-23.

Betz, C., and Hall, M.N. (2013). Where is mTOR and what is it doing there? *J Cell Biol* 203**,** 563-574.

Biederer, T., Sara, Y., Mozhayeva, M., Atasoy, D., Liu, X., Kavalali, E.T., and Sudhof, T.C. (2002). SynCAM, a synaptic adhesion molecule that drives synapse assembly. *Science* 297**,** 1525-1531.

Bigford, G.E., Alonso, O.F., Dietrich, D., and Keane, R.W. (2009). A novel protein complex in membrane rafts linking the NR2B glutamate receptor and autophagy is disrupted following traumatic brain injury. *J Neurotrauma* 26**,** 703-720.

Blackstone, C.D., Moss, S.J., Martin, L.J., Levey, A.I., Price, D.L., and Huganir, R.L. (1992). Biochemical characterization and localization of a non-N-methyl-D-aspartate glutamate receptor in rat brain. *J Neurochem* 58**,** 1118-1126.

Blanke, M.L., and Vandongen, A.M.J. (2009). "Activation Mechanisms of the NMDA Receptor," in *Biology of the NMDA Receptor,* ed. A.M. Van Dongen. (Boca Raton (FL)).

Blednov, Y.A., Benavidez, J.M., Black, M., Chandra, D., Homanics, G.E., Rudolph, U., and Harris, R.A. (2013). Linking GABA(A) receptor subunits to alcohol-induced conditioned taste aversion and recovery from acute alcohol intoxication. *Neuropharmacology* 67**,** 46-56.

Blednov, Y.A., Stoffel, M., Chang, S.R., and Harris, R.A. (2001). Potassium channels as targets for ethanol: studies of G-protein-coupled inwardly rectifying potassium channel 2 (GIRK2) null mutant mice. *J Pharmacol Exp Ther* 298**,** 521-530.

Blundell, J., Kouser, M., and Powell, C.M. (2008). Systemic inhibition of mammalian target of rapamycin inhibits fear memory reconsolidation. *Neurobiol Learn Mem* 90**,** 28-35.

Boden, J.M., and Fergusson, D.M. (2011). Alcohol and depression. *Addiction* 106**,** 906-914.

Bodian, D. (1965). A Suggestive Relationship of Nerve Cell Rna with Specific Synaptic Sites. *Proc Natl Acad Sci U S A* 53**,** 418-425.

Bodian, D. (1972). *Synaptic diversity and characterization by electron microscopy.* New York: Raven Press.

Bohnsack, M.T., Tollervey, D., and Granneman, S. (2012). Identification of RNA helicase target sites by UV cross-linking and analysis of cDNA. *Methods Enzymol* 511**,** 275-288.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30**,** 2114-2120.

Boratyn, G.M., Camacho, C., Cooper, P.S., Coulouris, G., Fong, A., Ma, N., Madden, T.L., Matten, W.T., Mcginnis, S.D., Merezhuk, Y., Raytselis, Y., Sayers, E.W., Tao, T., Ye, J., and Zaretskaya, I. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Res* 41**,** W29-33.

Boudreau, R.L., Jiang, P., Gilmore, B.L., Spengler, R.M., Tirabassi, R., Nelson, J.A., Ross, C.A., Xing, Y., and Davidson, B.L. (2014). Transcriptome-wide discovery of microRNA binding sites in human brain. *Neuron* 81**,** 294-305.

Bourgeron, T. (2009). A synaptic trek to autism. *Curr Opin Neurobiol* 19**,** 231-234.

Bourne, J.N., and Harris, K.M. (2008). Balancing structure and function at hippocampal dendritic spines. *Annu Rev Neurosci* 31**,** 47-67.

Bove, J., Martinez-Vicente, M., and Vila, M. (2011). Fighting neurodegeneration with rapamycin: mechanistic insights. *Nat Rev Neurosci* 12**,** 437-452.

Bowling, H., Bhattacharya, A., Zhang, G., Lebowitz, J.Z., Alam, D., Smith, P.T., Kirshenbaum, K., Neubert, T.A., Vogel, C., Chao, M.V., and Klann, E. (2016). BONLAC: A combinatorial proteomic technique to measure stimulus-induced translational profiles in brain slices. *Neuropharmacology* 100**,** 76-89.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol* 34**,** 525-527.

Brewster, A.L., Lugo, J.N., Patil, V.V., Lee, W.L., Qian, Y., Vanegas, F., and Anderson, A.E. (2013). Rapamycin reverses status epilepticus-induced memory deficits and dendritic damage. *PLoS One* 8**,** e57808.

Broek, J.A., Lin, Z., De Gruiter, H.M., Van 'T Spijker, H., Haasdijk, E.D., Cox, D., Ozcan, S., Van Cappellen, G.W., Houtsmuller, A.B., Willemsen, R., De Zeeuw, C.I., and Bahn, S. (2016). Synaptic vesicle dynamic changes in a model of fragile X. *Mol Autism* 7**,** 17.

Brosch, M., Yu, L., Hubbard, T., and Choudhary, J. (2009). Accurate and sensitive peptide identification with Mascot Percolator. *J Proteome Res* 8**,** 3176-3181.

Brown, V., Jin, P., Ceman, S., Darnell, J.C., O'donnell, W.T., Tenenbaum, S.A., Jin, X., Feng, Y., Wilkinson, K.D., Keene, J.D., Darnell, R.B., and Warren, S.T. (2001). Microarray identification of FMRP-associated brain mRNAs and altered mRNA translational profiles in fragile X syndrome. *Cell* 107**,** 477-487.

Buckanovich, R.J., and Darnell, R.B. (1997). The neuronal RNA binding protein Nova-1 recognizes specific RNA targets in vitro and in vivo. *Mol Cell Biol* 17**,** 3194-3201.

Buffington, S.A., Huang, W., and Costa-Mattioli, M. (2014). Translational control in synaptic plasticity and cognitive dysfunction. *Annu Rev Neurosci* 37**,** 17-38.

Buhr, F., Kohl-Landgraf, J., Tom Dieck, S., Hanus, C., Chatterjee, D., Hegelein, A., Schuman, E.M., Wachtveitl, J., and Schwalbe, H. (2015). Design of photocaged puromycin for nascent polypeptide release and spatiotemporal monitoring of translation. *Angew Chem Int Ed Engl* 54**,** 3717-3721.

Burger, K., Muhl, B., Kellner, M., Rohrmoser, M., Gruber-Eber, A., Windhager, L., Friedel, C.C., Dolken, L., and Eick, D. (2013). 4-thiouridine inhibits rRNA synthesis and causes a nucleolar stress response. *RNA Biol* 10**,** 1623-1630.

Burnashev, N., Schoepfer, R., Monyer, H., Ruppersberg, J.P., Gunther, W., Seeburg, P.H., and Sakmann, B. (1992). Control by asparagine residues of calcium permeability and magnesium blockade in the NMDA receptor. *Science* 257**,** 1415-1419.

Busby, M.A., Stewart, C., Miller, C.A., Grzeda, K.R., and Marth, G.T. (2013). Scotty: a web tool for designing RNA-Seq experiments to measure differential gene expression. *Bioinformatics* 29**,** 656-657.

Butter, F., Scheibe, M., Morl, M., and Mann, M. (2009). Unbiased RNA-protein interaction screen by quantitative proteomics. *Proc Natl Acad Sci U S A* 106**,** 10626-10631.

Butz, M., Worgotter, F., and Van Ooyen, A. (2009). Activity-dependent structural plasticity. *Brain Res Rev* 60**,** 287-305.

Buxbaum, A.R., Haimovich, G., and Singer, R.H. (2015a). In the right place at the right time: visualizing and understanding mRNA localization. *Nat Rev Mol Cell Biol* 16**,** 95-109.

Buxbaum, A.R., Yoon, Y.J., Singer, R.H., and Park, H.Y. (2015b). Single-molecule insights into mRNA dynamics in neurons. *Trends Cell Biol* 25**,** 468-475.

Cahoy, J.D., Emery, B., Kaushal, A., Foo, L.C., Zamanian, J.L., Christopherson, K.S., Xing, Y., Lubischer, J.L., Krieg, P.A., Krupenko, S.A., Thompson, W.J., and

Barres, B.A. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* 28**,** 264-278.

Cajigas, I.J., Tushev, G., Will, T.J., Tom Dieck, S., Fuerst, N., and Schuman, E.M. (2012). The local transcriptome in the synaptic neuropil revealed by deep sequencing and high-resolution imaging. *Neuron* 74**,** 453-466.

Cajigas, I.J., Will, T., and Schuman, E.M. (2010). Protein homeostasis and synaptic plasticity. *EMBO J* 29**,** 2746-2752.

Cammalleri, M., Lutjens, R., Berton, F., King, A.R., Simpson, C., Francesconi, W., and Sanna, P.P. (2003). Time-restricted role for dendritic activation of the mTOR-p70S6K pathway in the induction of late-phase long-term potentiation in the CA1. *Proc Natl Acad Sci U S A* 100**,** 14368-14373.

Campbell, Z.T., Bhimsaria, D., Valley, C.T., Rodriguez-Martinez, J.A., Menichelli, E., Williamson, J.R., Ansari, A.Z., and Wickens, M. (2012). Cooperativity in RNA-protein interactions: global analysis of RNA binding specificity. *Cell Rep* 1**,** 570-581.

Cao, Y., Charisi, A., Cheng, L.C., Jiang, T., and Girke, T. (2008). ChemmineR: a compound mining framework for R. *Bioinformatics* 24**,** 1733-1734.

Carlin, R.K., Grab, D.J., Cohen, R.S., and Siekevitz, P. (1980). Isolation and characterization of postsynaptic densities from various brain regions: enrichment of different types of postsynaptic densities. *J Cell Biol* 86**,** 831-845.

Carlson, M. (2016a). "mgu74a.db: Affymetrix Murine Genome U74v2 annotation data (chip mgu74a)". R package version 3.2.3 ed.).

Carlson, M. (2016b). "org.Hs.eg.db: Genome wide annotation for Human". R package version 3.4.0 ed.).

Carlson, M. (2016c). "org.Mm.eg.db: Genome wide annotation for Mouse". R package version 3.4.0 ed.).

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2013). System-wide identification of RNA-binding proteins by interactome capture. *Nat Protoc* 8**,** 491-500.

Castello, A., Horos, R., Strein, C., Fischer, B., Eichelbaum, K., Steinmetz, L.M., Krijgsveld, J., and Hentze, M.W. (2016). Comprehensive Identification of RNA-Binding Proteins by RNA Interactome Capture. *Methods Mol Biol* 1358**,** 131-139.

Catchen, J.M., Amores, A., Hohenlohe, P., Cresko, W., and Postlethwait, J.H. (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3 (Bethesda)* 1**,** 171-182.

Catterall, W.A., and Few, A.P. (2008). Calcium channel regulation and presynaptic plasticity. *Neuron* 59**,** 882-901.

Ch'ng, T.H., Desalvo, M., Lin, P., Vashisht, A., Wohlschlegel, J.A., and Martin, K.C. (2015). Cell biological mechanisms of activity-dependent synapse to nucleus translocation of CRTC1 in neurons. *Front Mol Neurosci* 8**,** 48.

Chandler, L.J., Sumners, C., and Crews, F.T. (1993). Ethanol inhibits NMDA receptor-mediated excitotoxicity in rat primary neuronal cultures. *Alcohol Clin Exp Res* 17**,** 54-60.

Chang, W., Cheng, J., Allaire, J.J., Xie, Y., and Mcpherson, J. (2016). "shiny: Web Application Framework for R". R package version 0.14.2 ed.).

Charif, D., and Lobry, J.R. (2007). "SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrival and analysis," in *Structural Approaches to Sequence Evolution: Molecules, Networks, Populations,* eds. U. Bastolla, M. Porto, H.E. Roman & M. Vendruscolo. (Berlin: Springer), 207-232.

Chen, X., Shu, S., and Bayliss, D.A. (2009). HCN1 channel subunits are a molecular substrate for hypnotic actions of ketamine. *J Neurosci* 29**,** 600-609.

Chen, X., Wei, S., Ji, Y., Guo, X., and Yang, F. (2015). Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* 15**,** 3175-3192.

Chernova, O.B., Somerville, R.P., and Cowell, J.K. (1998). A novel gene, LGI1, from 10q24 is rearranged and downregulated in malignant brain tumors. *Oncogene* 17**,** 2873-2881.

Chicurel, M.E., Terrian, D.M., and Potter, H. (Year). "Subcellular localization of mRNA: isolation and characterization of mRNA from an enriched preparation of hippocampal dendritic spines", in: *Society for Neuroscience*), 353.

Ching, T., Huang, S., and Garmire, L.X. (2014). Power analysis and sample size estimation for RNA-Seq differential expression. *RNA* 20**,** 1684-1696.

Christie, S.B., Akins, M.R., Schwob, J.E., and Fallon, J.R. (2009). The FXG: a presynaptic fragile X granule expressed in a subset of developing brain circuits. *J Neurosci* 29**,** 1514-1524.

Chudakov, D.M., Lukyanov, S., and Lukyanov, K.A. (2007). Tracking intracellular protein movements using photoswitchable fluorescent proteins PS-CFP2 and Dendra2. *Nat Protoc* 2**,** 2024-2032.

Chudakov, D.M., Matz, M.V., Lukyanov, S., and Lukyanov, K.A. (2010). Fluorescent proteins and their applications in imaging living cells and tissues. *Physiol Rev* 90**,** 1103-1163.

Cieslik, M., Chugh, R., Wu, Y.M., Wu, M., Brennan, C., Lonigro, R., Su, F., Wang, R., Siddiqui, J., Mehra, R., Cao, X., Lucas, D., Chinnaiyan, A.M., and Robinson, D. (2015). The use of exome capture RNA-seq for highly degraded RNA with application to clinical cancer sequencing. *Genome Res* 25**,** 1372-1381.

Cirelli, C., Bushey, D., Hill, S., Huber, R., Kreber, R., Ganetzky, B., and Tononi, G. (2005). Reduced sleep in Drosophila Shaker mutants. *Nature* 434**,** 1087-1092.

Cohen, R.S., Blomberg, F., Berzins, K., and Siekevitz, P. (1977). The structure of postsynaptic densities isolated from dog cerebral cortex. I. Overall morphology and protein composition. *J Cell Biol* 74**,** 181-203.

Collins, M.O., Husi, H., Yu, L., Brandon, J.M., Anderson, C.N., Blackstock, W.P., Choudhary, J.S., and Grant, S.G. (2006). Molecular characterization and

comparison of the components and multiprotein complexes in the postsynaptic proteome. *J Neurochem* 97 Suppl 1**,** 16-23.

Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., Mcpherson, A., Szczesniak, M.W., Gaffney, D.J., Elo, L.L., Zhang, X., and Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biol* 17**,** 13.

Cookson, M.R. (2004). Molecules that cause or prevent Parkinson's disease. *PLoS Biol* 2**,** e401.

Costa-Mattioli, M., and Monteggia, L.M. (2013). mTOR complexes in neurodevelopmental and neuropsychiatric disorders. *Nat Neurosci* 16**,** 1537-1543.

Costa-Mattioli, M., Sossin, W.S., Klann, E., and Sonenberg, N. (2009). Translational control of long-lasting synaptic plasticity and memory. *Neuron* 61**,** 10-26.

Crino, P.B. (2008). Do we have a cure for tuberous sclerosis complex? *Epilepsy Curr* 8**,** 159-162.

Crino, P.B. (2015). mTOR signaling in epilepsy: insights from malformations of cortical development. *Cold Spring Harb Perspect Med* 5.

Cui, P., Lin, Q., Ding, F., Xin, C., Gong, W., Zhang, L., Geng, J., Zhang, B., Yu, X., Yang, J., Hu, S., and Yu, J. (2010). A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. *Genomics* 96**,** 259-265.

Cui, X. (2010). "Experimental designs on high-throughput biological experiments," in *Statistical Bioinformatics*. (New Jersey: Wiley-Blackwell).

Cunningham, F., Amode, M.R., Barrell, D., Beal, K., Billis, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fitzgerald, S., Gil, L., Giron, C.G., Gordon, L., Hourlier, T., Hunt, S.E., Janacek, S.H., Johnson, N., Juettemann, T., Kahari, A.K., Keenan, S., Martin, F.J., Maurel, T., Mclaren, W., Murphy, D.N., Nag, R., Overduin, B., Parker, A., Patricio, M., Perry, E., Pignatelli, M., Riat, H.S., Sheppard, D., Taylor, K., Thormann, A., Vullo, A., Wilder, S.P., Zadissa, A., Aken, B.L., Birney, E., Harrow, J., Kinsella, R., Muffato, M., Ruffier, M., Searle, S.M., Spudich, G., Trevanion, S.J., Yates, A., Zerbino, D.R., and Flicek, P. (2015). Ensembl 2015. *Nucleic Acids Res* 43**,** D662-669.

D'haeseleer, P. (2005). How does gene expression clustering work? *Nat Biotechnol* 23**,** 1499-1501.

D'hulst, C., De Geest, N., Reeve, S.P., Van Dam, D., De Deyn, P.P., Hassan, B.A., and Kooy, R.F. (2006). Decreased expression of the GABAA receptor in fragile X syndrome. *Brain Res* 1121**,** 238-245.

Dahm, G.M., Gubin, M.M., Magee, J.D., Techasintana, P., Calaluce, R., and Atasoy, U. (2012). Method for the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts using RIP-Chip. *J Vis Exp*.

Darmanis, S., Sloan, S.A., Zhang, Y., Enge, M., Caneda, C., Shuer, L.M., Hayden Gephart, M.G., Barres, B.A., and Quake, S.R. (2015). A survey of human brain transcriptome diversity at the single cell level. *Proc Natl Acad Sci U S A* 112**,** 7285-7290.

Darmostuk, M., Rimpelova, S., Gbelcova, H., and Ruml, T. (2015). Current approaches in SELEX: An update to aptamer selection technology. *Biotechnol Adv* 33**,** 1141-1161.

Darnell, J.C., Van Driesche, S.J., Zhang, C., Hung, K.Y., Mele, A., Fraser, C.E., Stone, E.F., Chen, C., Fak, J.J., Chi, S.W., Licatalosi, D.D., Richter, J.D., and Darnell, R.B. (2011). FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* 146**,** 247-261.

Darnell, R.B. (2010). HITS-CLIP: panoramic views of protein-RNA regulation in living cells. *Wiley Interdiscip Rev RNA* 1**,** 266-286.

David, D.J., Samuels, B.A., Rainer, Q., Wang, J.W., Marsteller, D., Mendez, I., Drew, M., Craig, D.A., Guiard, B.P., Guilloux, J.P., Artymyshyn, R.P., Gardier, A.M., Gerald, C., Antonijevic, I.A., Leonardo, E.D., and Hen, R. (2009). Neurogenesis-dependent and -independent effects of fluoxetine in an animal model of anxiety/depression. *Neuron* 62**,** 479-493.

David, M., Dzamba, M., Lister, D., Ilie, L., and Brudno, M. (2011). SHRiMP2: sensitive yet practical SHort Read Mapping. *Bioinformatics* 27**,** 1011-1012.

Davis, G.W., and Goodman, C.S. (1998). Genetic analysis of synaptic development and plasticity: homeostatic regulation of synaptic efficacy. *Curr Opin Neurobiol* 8**,** 149-156.

Dazert, E., and Hall, M.N. (2011). mTOR signaling in disease. *Curr Opin Cell Biol* 23**,** 744-755.

De Bartolomeis, A., Sarappa, C., Buonaguro, E.F., Marmo, F., Eramo, A., Tomasetti, C., and Iasevoli, F. (2013). Different effects of the NMDA receptor antagonists ketamine, MK-801, and memantine on postsynaptic density transcripts and their topography: role of Homer signaling, and implications for novel antipsychotic and pro-cognitive targets in psychosis. *Prog Neuropsychopharmacol Biol Psychiatry* 46**,** 1-12.

De Rubeis, S., Fernandez, E., Buzzi, A., Di Marino, D., and Bagni, C. (2012). Molecular and cellular aspects of mental retardation in the Fragile X syndrome: from gene mutation/s to spine dysmorphogenesis. *Adv Exp Med Biol* 970**,** 517-551.

De Vos, S., Wardenaar, K.J., Bos, E.H., Wit, E.C., and De Jonge, P. (2015). Decomposing the heterogeneity of depression at the person-, symptom-, and time-level: latent variable models versus multimode principal component analysis. *BMC Med Res Methodol* 15**,** 88.

De Vry, J., Vanmierlo, T., Martinez-Martinez, P., Losen, M., Temel, Y., Boere, J., Kenis, G., Steckler, T., Steinbusch, H.W., De Baets, M., and Prickaerts, J. (2016). TrkB in the hippocampus and nucleus accumbens differentially modulates depression-like behavior in mice. *Behav Brain Res* 296**,** 15-25.

Dekker, M.J., Tiemeier, H., Luijendijk, H.J., Kuningas, M., Hofman, A., De Jong, F.H., Stewart, P.M., Koper, J.W., and Lamberts, S.W. (2012). The effect of common genetic variation in 11beta-hydroxysteroid dehydrogenase type 1 on hypothalamic-pituitary-adrenal axis activity and incident depression. *J Clin Endocrinol Metab* 97**,** E233-237.

Deluca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* 28**,** 1530-1532.

Deschwanden, A., Karolewicz, B., Feyissa, A.M., Treyer, V., Ametamey, S.M., Johayem, A., Burger, C., Auberson, Y.P., Sovago, J., Stockmeier, C.A., Buck, A., and Hasler, G. (2011). Reduced metabotropic glutamate receptor 5 density in major depression determined by [(11)C]ABP688 PET and postmortem study. *Am J Psychiatry* 168**,** 727-734.

Devi, L., and Ohno, M. (2013). Deletion of the eIF2alpha Kinase GCN2 fails to rescue the memory decline associated with Alzheimer's disease. *PLoS One* 8**,** e77335.

Di Niro, R., Sulic, A.M., Mignone, F., D'angelo, S., Bordoni, R., Iacono, M., Marzari, R., Gaiotto, T., Lavric, M., Bradbury, A.R., Biancone, L., Zevin-Sonkin, D., De Bellis, G., Santoro, C., and Sblattero, D. (2010). Rapid interactome profiling by massive sequencing. *Nucleic Acids Res* 38**,** e110.

Dictenberg, J.B., Swanger, S.A., Antar, L.N., Singer, R.H., and Bassell, G.J. (2008). A direct role for FMRP in activity-dependent dendritic mRNA transport links filopodial-spine morphogenesis to fragile X syndrome. *Dev Cell* 14**,** 926-939.

Dieterich, D.C., Hodas, J.J., Gouzer, G., Shadrin, I.Y., Ngo, J.T., Triller, A., Tirrell, D.A., and Schuman, E.M. (2010). In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci* 13**,** 897-905.

Dieterich, D.C., Lee, J.J., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2007). Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging. *Nat Protoc* 2**,** 532-540.

Dieterich, D.C., Link, A.J., Graumann, J., Tirrell, D.A., and Schuman, E.M. (2006). Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A* 103**,** 9482-9487.

Dillies, M.A., Rau, A., Aubert, J., Hennequet-Antier, C., Jeanmougin, M., Servant, N., Keime, C., Marot, G., Castel, D., Estelle, J., Guernec, G., Jagla, B., Jouneau, L., Laloe, D., Le Gall, C., Schaeffer, B., Le Crom, S., Guedj, M., Jaffrezic, F., and French Statomique, C. (2013). A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. *Brief Bioinform* 14**,** 671-683.

Dlugos, A.M., Palmer, A.A., and De Wit, H. (2009). Negative emotionality: monoamine oxidase B gene variants modulate personality traits in healthy humans. *J Neural Transm (Vienna)* 116**,** 1323-1334.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29**,** 15-21.

Donnelly, C.J., Fainzilber, M., and Twiss, J.L. (2010). Subcellular communication through RNA transport and localized protein synthesis. *Traffic* 11**,** 1498-1505.

Doyle, M., and Kiebler, M.A. (2011). Mechanisms of dendritic mRNA transport and its role in synaptic tagging. *EMBO J* 30**,** 3540-3552.

Dragulescu, A.A. (2014). "xlsx: Read, write, format Excel 2007 and Excel 97/2000/XP/2003 files". R package version 0.5.7 ed.).

Dudchenko, P.A. (2004). An overview of the tasks used to test working memory in rodents. *Neurosci Biobehav Rev* 28**,** 699-709.

Dueck, H., Eberwine, J., and Kim, J. (2016). Variation is function: Are single cell differences functionally important?: Testing the hypothesis that single cell variation is required for aggregate function. *Bioessays* 38**,** 172-180.

Dueck, H., Khaladkar, M., Kim, T.K., Spaethling, J.M., Francis, C., Suresh, S., Fisher, S.A., Seale, P., Beck, S.G., Bartfai, T., Kuhn, B., Eberwine, J., and Kim, J. (2015). Deep sequencing reveals cell-type-specific patterns of single-cell transcriptome variation. *Genome Biol* 16**,** 122.

Durand, C.M., Perroy, J., Loll, F., Perrais, D., Fagni, L., Bourgeron, T., Montcouquiol, M., and Sans, N. (2012). SHANK3 mutations identified in autism lead to modification of dendritic spine morphology via an actin-dependent mechanism. *Mol Psychiatry* 17**,** 71-84.

Durinck, S., Moreau, Y., Kasprzyk, A., Davis, S., De Moor, B., Brazma, A., and Huber, W. (2005). BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics* 21**,** 3439-3440.

Durinck, S., Spellman, P.T., Birney, E., and Huber, W. (2009). Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 4**,** 1184-1191.

Dwyer, J.M., and Duman, R.S. (2013). Activation of mammalian target of rapamycin and synaptogenesis: role in the actions of rapid-acting antidepressants. *Biol Psychiatry* 73**,** 1189-1198.

Eberwine, J., Miyashiro, K., Kacharmina, J.E., and Job, C. (2001). Local translation of classes of mRNAs that are targeted to neuronal dendrites. *Proc Natl Acad Sci U S A* 98**,** 7080-7085.

Ekstrand, M.I., Nectow, A.R., Knight, Z.A., Latcha, K.N., Pomeranz, L.E., and Friedman, J.M. (2014). Molecular profiling of neurons based on connectivity. *Cell* 157**,** 1230-1242.

Engstrom, P.G., Steijger, T., Sipos, B., Grant, G.R., Kahles, A., Ratsch, G., Goldman, N., Hubbard, T.J., Harrow, J., Guigo, R., Bertone, P., and Consortium, R. (2013). Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat Methods* 10**,** 1185-1191.

Erhard, F., Dolken, L., and Zimmer, R. (2013). RIP-chip enrichment analysis. *Bioinformatics* 29**,** 77-83.

Esposito, A.M., Mateyak, M., He, D., Lewis, M., Sasikumar, A.N., Hutton, J., Copeland, P.R., and Kinzy, T.G. (2010). Eukaryotic polyribosome profile analysis. *J Vis Exp*.

Fang, Z., and Cui, X. (2011). Design and validation issues in RNA-seq experiments. *Brief Bioinform* 12**,** 280-287.

Farris, S.P., Arasappan, D., Hunicke-Smith, S., Harris, R.A., and Mayfield, R.D. (2015). Transcriptome organization for chronic alcohol abuse in human brain. *Mol Psychiatry* 20**,** 1438-1447.

Fatemi, S.H., Earle, J.A., Stary, J.M., Lee, S., and Sedgewick, J. (2001). Altered levels of the synaptosomal associated protein SNAP-25 in hippocampus of subjects with mood disorders and schizophrenia. *Neuroreport* 12**,** 3257-3262.

Feng, Z., Zhang, H., Levine, A.J., and Jin, S. (2005). The coordinate regulation of the p53 and mTOR pathways in cells. *Proc Natl Acad Sci U S A* 102**,** 8204-8209.

Feret, J., Danos, V., Krivine, J., Harmer, R., and Fontana, W. (2009). Internal coarse-graining of molecular systems. *Proc Natl Acad Sci U S A* 106**,** 6453-6458.

Fernandez-Moya, S.M., Bauer, K.E., and Kiebler, M.A. (2014). Meet the players: local translation at the synapse. *Front Mol Neurosci* 7**,** 84.

Fernandez, E., Collins, M.O., Uren, R.T., Kopanitsa, M.V., Komiyama, N.H., Croning, M.D., Zografos, L., Armstrong, J.D., Choudhary, J.S., and Grant, S.G. (2009). Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins. *Mol Syst Biol* 5**,** 269.

Fernandez, E., Li, K.W., Rajan, N., De Rubeis, S., Fiers, M., Smit, A.B., Achsel, T., and Bagni, C. (2015). FXR2P Exerts a Positive Translational Control and Is Required for the Activity-Dependent Increase of PSD95 Expression. *J Neurosci* 35**,** 9402-9408.

Feyissa, A.M., Chandran, A., Stockmeier, C.A., and Karolewicz, B. (2009). Reduced levels of NR2A and NR2B subunits of NMDA receptor and PSD-95 in the prefrontal cortex in major depression. *Prog Neuropsychopharmacol Biol Psychiatry* 33**,** 70-75.

Finotello, F., and Di Camillo, B. (2015). Measuring differential gene expression with RNA-seq: challenges and strategies for data analysis. *Brief Funct Genomics* 14**,** 130-142.

Fiszer, S., and Robertis, E.D. (1967). Action of triton X-100 on ultrastructure and membrane-bound- enzymes of isolated nerve endings from rat brain. *Brain Res* 5**,** 31-44.

Flynn, J.M., Czerwieniec, G.A., Choi, S.W., Day, N.U., Gibson, B.W., Hubbard, A., and Melov, S. (2012). Proteogenomics of synaptosomal mitochondrial oxidative stress. *Free Radic Biol Med* 53**,** 1048-1060.

Folsom, T.D., Thuras, P.D., and Fatemi, S.H. (2015). Protein expression of targets of the FMRP regulon is altered in brains of subjects with schizophrenia and mood disorders. *Schizophr Res* 165**,** 201-211.

Fonseca, N.A., Rung, J., Brazma, A., and Marioni, J.C. (2012). Tools for mapping high-throughput sequencing data. *Bioinformatics* 28**,** 3169-3177.

Fossati, G., Morini, R., Corradini, I., Antonucci, F., Trepte, P., Edry, E., Sharma, V., Papale, A., Pozzi, D., Defilippi, P., Meier, J.C., Brambilla, R., Turco, E., Rosenblum, K., Wanker, E.E., Ziv, N.E., Menna, E., and Matteoli, M. (2015). Reduced SNAP-25 increases PSD-95 mobility and impairs spine morphogenesis. *Cell Death Differ* 22**,** 1425-1436.

Fournier, J.C., Derubeis, R.J., Hollon, S.D., Dimidjian, S., Amsterdam, J.D., Shelton, R.C., and Fawcett, J. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* 303**,** 47-53.

Fragkouli, A., and Doxakis, E. (2014). miR-7 and miR-153 protect neurons against MPP(+)-induced cell death via upregulation of mTOR pathway. *Front Cell Neurosci* 8**,** 182.

Free, R.B., Hazelwood, L.A., and Sibley, D.R. (2009). Identifying novel protein-protein interactions using co-immunoprecipitation and mass spectroscopy. *Curr Protoc Neurosci* Chapter 5**,** Unit 5 28.

Fresno, C., and Fernandez, E.A. (2013). RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics* 29**,** 2810-2811.

Frey, U., and Morris, R.G. (1997). Synaptic tagging and long-term potentiation. *Nature* 385**,** 533-536.

Fried, E.I. (2017). The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. *J Affect Disord* 208**,** 191-197.

Friedersdorf, M.B., and Keene, J.D. (2014). Advancing the functional utility of PAR-CLIP by quantifying background binding to mRNAs and lncRNAs. *Genome Biol* 15**,** R2.

Fritzsche, R., Karra, D., Bennett, K.L., Ang, F.Y., Heraud-Farlow, J.E., Tolino, M., Doyle, M., Bauer, K.E., Thomas, S., Planyavsky, M., Arn, E., Bakosova, A., Jungwirth, K., Hormann, A., Palfi, Z., Sandholzer, J., Schwarz, M., Macchi, P., Colinge, J., Superti-Furga, G., and Kiebler, M.A. (2013). Interactome of two diverse RNA granules links mRNA localization to translational repression in neurons. *Cell Rep* 5**,** 1749-1762.

Fu, X., Fu, N., Guo, S., Yan, Z., Xu, Y., Hu, H., Menzel, C., Chen, W., Li, Y., Zeng, R., and Khaitovich, P. (2009). Estimating accuracy of RNA-Seq and microarrays with proteomics. *BMC Genomics* 10**,** 161.

Fuchsova, B., Alvarez Julia, A., Rizavi, H.S., Frasch, A.C., and Pandey, G.N. (2015). Altered expression of neuroplasticity-related genes in the brain of depressed suicides. *Neuroscience* 299**,** 1-17.

Fujii, R., and Takumi, T. (2005). TLS facilitates transport of mRNA encoding an actin-stabilizing protein to dendritic spines. *J Cell Sci* 118**,** 5755-5765.

Fukata, Y., Adesnik, H., Iwanaga, T., Bredt, D.S., Nicoll, R.A., and Fukata, M. (2006). Epilepsy-related ligand/receptor complex LGI1 and ADAM22 regulate synaptic transmission. *Science* 313**,** 1792-1795.

Fukata, Y., Lovero, K.L., Iwanaga, T., Watanabe, A., Yokoi, N., Tabuchi, K., Shigemoto, R., Nicoll, R.A., and Fukata, M. (2010). Disruption of LGI1-linked synaptic complex causes abnormal synaptic transmission and epilepsy. *Proc Natl Acad Sci U S A* 107**,** 3799-3804.

Fuller, J.C., Martinez, M., Henrich, S., Stank, A., Richter, S., and Wade, R.C. (2015). LigDig: a web server for querying ligand-protein interactions. *Bioinformatics* 31**,** 1147-1149.

Gafford, G.M., Parsons, R.G., and Helmstetter, F.J. (2011). Consolidation and reconsolidation of contextual fear memory requires mammalian target of rapamycin-dependent translation in the dorsal hippocampus. *Neuroscience* 182**,** 98-104.

Galeotti, N., and Ghelardini, C. (2011). Antidepressant phenotype by inhibiting the phospholipase Cbeta(1)--protein kinase Cgamma pathway in the forced swim test. *Neuropharmacology* 60**,** 937-943.

Gallego Romero, I., Pai, A.A., Tung, J., and Gilad, Y. (2014). RNA-seq: impact of RNA degradation on transcript quantification. *BMC Biol* 12**,** 42.

Gandin, V., Sikstrom, K., Alain, T., Morita, M., Mclaughlan, S., Larsson, O., and Topisirovic, I. (2014). Polysome fractionation and analysis of mammalian translatomes on a genome-wide scale. *J Vis Exp*.

Ganley, I.G., Lam Du, H., Wang, J., Ding, X., Chen, S., and Jiang, X. (2009). ULK1.ATG13.FIP200 complex mediates mTOR signaling and is essential for autophagy. *J Biol Chem* 284**,** 12297-12305.

Gatt, J.M., Burton, K.L., Williams, L.M., and Schofield, P.R. (2015). Specific and common genes implicated across major mental disorders: a review of meta-analysis studies. *J Psychiatr Res* 60**,** 1-13.

Gatto, C.L., Pereira, D., and Broadie, K. (2014). GABAergic circuit dysfunction in the Drosophila Fragile X syndrome model. *Neurobiol Dis* 65**,** 142-159.

Genheden, M., Kenney, J.W., Johnston, H.E., Manousopoulou, A., Garbis, S.D., and Proud, C.G. (2015). BDNF stimulation of protein synthesis in cortical neurons requires the MAP kinase-interacting kinase MNK1. *J Neurosci* 35**,** 972-984.

Geschwind, D.H., and Konopka, G. (2009). Neuroscience in the era of functional genomics and systems biology. *Nature* 461**,** 908-915.

Gibson, J.R., Bartley, A.F., Hays, S.A., and Huber, K.M. (2008). Imbalance of neocortical excitation and inhibition and altered UP states reflect network hyperexcitability in the mouse model of fragile X syndrome. *J Neurophysiol* 100**,** 2615-2626.

Gideons, E.S., Kavalali, E.T., and Monteggia, L.M. (2014). Mechanisms underlying differential effectiveness of memantine and ketamine in rapid antidepressant responses. *Proc Natl Acad Sci U S A* 111**,** 8649-8654.

Giese, K.P. (2012). "Long-term potentiation and memory," in *Memory mechanisms in health and disease* (Singapore: World Scientific Publishing), Giese, K.P.

Giese, K.P., and Mizuno, K. (2013). The roles of protein kinases in learning and memory. *Learn Mem* 20**,** 540-552.

Gillen, A.E., Yamamoto, T.M., Kline, E., Hesselberth, J.R., and Kabos, P. (2016). Improvements to the HITS-CLIP protocol eliminate widespread mispriming artifacts. *BMC Genomics* 17**,** 338.

Gingras, A.C., Kennedy, S.G., O'leary, M.A., Sonenberg, N., and Hay, N. (1998). 4E-BP1, a repressor of mRNA translation, is phosphorylated and inactivated by the Akt(PKB) signaling pathway. *Genes Dev* 12**,** 502-513.

Gkogkas, C.G., Khoutorsky, A., Ran, I., Rampakakis, E., Nevarko, T., Weatherill, D.B., Vasuta, C., Yee, S., Truitt, M., Dallaire, P., Major, F., Lasko, P., Ruggero, D., Nader, K., Lacaille, J.C., and Sonenberg, N. (2013). Autism-related deficits via dysregulated eIF4E-dependent translational control. *Nature* 493**,** 371-377.

Glasgow, N.G., Siegler Retchless, B., and Johnson, J.W. (2015). Molecular bases of NMDA receptor subtype-dependent properties. *J Physiol* 593**,** 83-95.

Goetze, B., Tuebing, F., Xie, Y., Dorostkar, M.M., Thomas, S., Pehl, U., Boehm, S., Macchi, P., and Kiebler, M.A. (2006). The brain-specific double-stranded RNA-binding protein Staufen2 is required for dendritic spine morphogenesis. *J Cell Biol* 172**,** 221-231.

Goldberg, D. (2011). The heterogeneity of "major depression". *World Psychiatry* 10**,** 226-228.

Gong, R., Park, C.S., Abbassi, N.R., and Tang, S.J. (2006). Roles of glutamate receptors and the mammalian target of rapamycin (mTOR) signaling pathway in activity-dependent dendritic protein synthesis in hippocampal neurons. *J Biol Chem* 281**,** 18802-18815.

Goodwin, S., Mcpherson, J.D., and Mccombie, W.R. (2016). Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 17**,** 333-351.

Graber, T.E., Hebert-Seropian, S., Khoutorsky, A., David, A., Yewdell, J.W., Lacaille, J.C., and Sossin, W.S. (2013a). Reactivation of stalled polyribosomes in synaptic plasticity. *Proc Natl Acad Sci U S A* 110**,** 16205-16210.

Graber, T.E., Mccamphill, P.K., and Sossin, W.S. (2013b). A recollection of mTOR signaling in learning and memory. *Learn Mem* 20**,** 518-530.

Granneman, S., Kudla, G., Petfalski, E., and Tollervey, D. (2009). Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc Natl Acad Sci U S A* 106**,** 9613-9618.

Gray, A.L., Hyde, T.M., Deep-Soboslay, A., Kleinman, J.E., and Sodhi, M.S. (2015). Sex differences in glutamate receptor gene expression in major depression and suicide. *Mol Psychiatry* 20**,** 1139.

Greber, S., Lubec, G., Cairns, N., and Fountoulakis, M. (1999). Decreased levels of synaptosomal associated protein 25 in the brain of patients with Down syndrome and Alzheimer's disease. *Electrophoresis* 20**,** 928-934.

Greenberg, P.E., Fournier, A.A., Sisitsky, T., Pike, C.T., and Kessler, R.C. (2015). The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psychiatry* 76**,** 155-162.

Gross, C., Nakamoto, M., Yao, X., Chan, C.B., Yim, S.Y., Ye, K., Warren, S.T., and Bassell, G.J. (2010). Excess phosphoinositide 3-kinase subunit synthesis and activity as a novel therapeutic target in fragile X syndrome. *J Neurosci* 30**,** 10624-10638.

Guetg, N., Abdel Aziz, S., Holbro, N., Turecek, R., Rose, T., Seddik, R., Gassmann, M., Moes, S., Jenoe, P., Oertner, T.G., Casanova, E., and Bettler, B. (2010). NMDA receptor-dependent GABAB receptor internalization via CaMKII phosphorylation of serine 867 in GABAB1. *Proc Natl Acad Sci U S A* 107**,** 13924-13929.

Gurskaya, N.G., Verkhusha, V.V., Shcheglov, A.S., Staroverov, D.B., Chepurnykh, T.V., Fradkov, A.F., Lukyanov, S., and Lukyanov, K.A. (2006). Engineering of a monomeric green-to-red photoactivatable fluorescent protein induced by blue light. *Nat Biotechnol* 24**,** 461-465.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jr., Jungkamp, A.C., Munschauer, M., Ulrich, A.,

Wardle, G.S., Dewell, S., Zavolan, M., and Tuschl, T. (2010). Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 141**,** 129-141.

Hafner, M., Lianoglou, S., Tuschl, T., and Betel, D. (2012). Genome-wide identification of miRNA targets by PAR-CLIP. *Methods* 58**,** 94-105.

Hagerman, R., and Hagerman, P. (2013). Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome. *Lancet Neurol* 12**,** 786-798.

Haider, S., and Pal, R. (2013). Integrated analysis of transcriptomic and proteomic data. *Curr Genomics* 14**,** 91-110.

Han, M.H., Jiao, S., Jia, J.M., Chen, Y., Chen, C.Y., Gucek, M., Markey, S.P., and Li, Z. (2013). The novel caspase-3 substrate Gap43 is involved in AMPA receptor endocytosis and long-term depression. *Mol Cell Proteomics* 12**,** 3719-3731.

Hanson, J.E., Weber, M., Meilandt, W.J., Wu, T., Luu, T., Deng, L., Shamloo, M., Sheng, M., Scearce-Levie, K., and Zhou, Q. (2013). GluN2B antagonism affects interneurons and leads to immediate and persistent changes in synaptic plasticity, oscillations, and behavior. *Neuropsychopharmacology* 38**,** 1221-1233.

Hara, K., Maruki, Y., Long, X., Yoshino, K., Oshiro, N., Hidayat, S., Tokunaga, C., Avruch, J., and Yonezawa, K. (2002). Raptor, a binding partner of target of rapamycin (TOR), mediates TOR action. *Cell* 110**,** 177-189.

Harbom, L.J., Chronister, W.D., and Mcconnell, M.J. (2016). Single neuron transcriptome analysis can reveal more than cell type classification: Does it matter if every neuron is unique? *Bioessays* 38**,** 157-161.

Hardcastle, T.J., and Kelly, K.A. (2010). baySeq: empirical Bayesian methods for identifying differential expression in sequence count data. *BMC Bioinformatics* 11**,** 422.

Hardie, D.G. (2007). AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nat Rev Mol Cell Biol* 8**,** 774-785.

Hart, S.N., Therneau, T.M., Zhang, Y., Poland, G.A., and Kocher, J.P. (2013). Calculating sample size estimates for RNA sequencing data. *J Comput Biol* 20**,** 970-978.

Hartl, D., Irmler, M., Romer, I., Mader, M.T., Mao, L., Zabel, C., De Angelis, M.H., Beckers, J., and Klose, J. (2008). Transcriptome and proteome analysis of early embryonic mouse brain development. *Proteomics* 8**,** 1257-1265.

Hay, N., and Sonenberg, N. (2004). Upstream and downstream of mTOR. *Genes Dev* 18**,** 1926-1945.

Hayashi, M.K., Ames, H.M., and Hayashi, Y. (2006). Tetrameric hub structure of postsynaptic scaffolding protein homer. *J Neurosci* 26**,** 8492-8501.

Hays, S.A., Huber, K.M., and Gibson, J.R. (2011). Altered neocortical rhythmic activity states in Fmr1 KO mice are due to enhanced mGluR5 signaling and involve changes in excitatory circuitry. *J Neurosci* 31**,** 14223-14234.

Head, S.R., Komori, H.K., Lamere, S.A., Whisenant, T., Van Nieuwerburgh, F., Salomon, D.R., and Ordoukhanian, P. (2014). Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* 56**,** 61-64, 66, 68, passim.

Hebb, C.O., and Smallman, B.N. (1956). Intracellular distribution of choline acetylase. *J Physiol* 134**,** 385-392.

Heiman, M., Schaefer, A., Gong, S., Peterson, J.D., Day, M., Ramsey, K.E., Suarez-Farinas, M., Schwarz, C., Stephan, D.A., Surmeier, D.J., Greengard, P., and Heintz, N. (2008). A translational profiling approach for the molecular characterization of CNS cell types. *Cell* 135**,** 738-748.

Hendel, T., Mank, M., Schnell, B., Griesbeck, O., Borst, A., and Reiff, D.F. (2008). Fluorescence changes of genetic calcium indicators and OGB-1 correlated with neural activity and calcium in vivo and in vitro. *J Neurosci* 28**,** 7399-7411.

Henderson, C., Wijetunge, L., Kinoshita, M.N., Shumway, M., Hammond, R.S., Postma, F.R., Brynczka, C., Rush, R., Thomas, A., Paylor, R., Warren, S.T., Vanderklish, P.W., Kind, P.C., Carpenter, R.L., Bear, M.F., and Healy, A.M. (2012). Reversal of disease-related pathologies in the fragile X mouse model by selective activation of GABAB receptors with arbaclofen. *Sci Transl Med* 4**,** 152ra128.

Henry, F.E., Mccartney, A.J., Neely, R., Perez, A.S., Carruthers, C.J., Stuenkel, E.L., Inoki, K., and Sutton, M.A. (2012). Retrograde changes in presynaptic function driven by dendritic mTORC1. *J Neurosci* 32**,** 17128-17142.

Hinds, H.L., Ashley, C.T., Sutcliffe, J.S., Nelson, D.L., Warren, S.T., Housman, D.E., and Schalling, M. (1993). Tissue specific expression of FMR-1 provides evidence for a functional role in fragile X syndrome. *Nat Genet* 3**,** 36-43.

Hoeffer, C.A., and Klann, E. (2010). mTOR signaling: at the crossroads of plasticity, memory and disease. *Trends Neurosci* 33**,** 67-75.

Hoeffer, C.A., Sanchez, E., Hagerman, R.J., Mu, Y., Nguyen, D.V., Wong, H., Whelan, A.M., Zukin, R.S., Klann, E., and Tassone, F. (2012). Altered mTOR signaling and enhanced CYFIP2 expression levels in subjects with fragile X syndrome. *Genes Brain Behav* 11**,** 332-341.

Hoell, J.I., Larsson, E., Runge, S., Nusbaum, J.D., Duggimpudi, S., Farazi, T.A., Hafner, M., Borkhardt, A., Sander, C., and Tuschl, T. (2011). RNA targets of wild-type and mutant FET family proteins. *Nat Struct Mol Biol* 18**,** 1428-1431.

Holahan, M.R., Honegger, K.S., Tabatadze, N., and Routtenberg, A. (2007). GAP-43 gene expression regulates information storage. *Learn Mem* 14**,** 407-415.

Hollingsworth, E.B., Mcneal, E.T., Burton, J.L., Williams, R.J., Daly, J.W., and Creveling, C.R. (1985). Biochemical characterization of a filtered synaptoneurosome preparation from guinea pig cerebral cortex: cyclic adenosine 3':5'-monophosphate-generating systems, receptors, and enzymes. *J Neurosci* 5**,** 2240-2253.

Holmes, A., Spanagel, R., and Krystal, J.H. (2013). Glutamatergic targets for new alcohol medications. *Psychopharmacology (Berl)* 229**,** 539-554.

Honer, W.G. (2003). Pathology of presynaptic proteins in Alzheimer's disease: more than simple loss of terminals. *Neurobiol Aging* 24**,** 1047-1062.

Honer, W.G., Barr, A.M., Sawada, K., Thornton, A.E., Morris, M.C., Leurgans, S.E., Schneider, J.A., and Bennett, D.A. (2012). Cognitive reserve, presynaptic proteins and dementia in the elderly. *Transl Psychiatry* 2**,** e114.

Hou, L., Antion, M.D., Hu, D., Spencer, C.M., Paylor, R., and Klann, E. (2006). Dynamic translational and proteasomal regulation of fragile X mental retardation protein controls mGluR-dependent long-term depression. *Neuron* 51**,** 441-454.

Hou, L., and Klann, E. (2004). Activation of the phosphoinositide 3-kinase-Akt-mammalian target of rapamycin signaling pathway is required for metabotropic glutamate receptor-dependent long-term depression. *J Neurosci* 24**,** 6352-6361.

Huang Da, W., Sherman, B.T., and Lempicki, R.A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 37**,** 1-13.

Huang Da, W., Sherman, B.T., and Lempicki, R.A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4**,** 44-57.

Huang, J., and Manning, B.D. (2008). The TSC1-TSC2 complex: a molecular switchboard controlling cell growth. *Biochem J* 412**,** 179-190.

Huber, K.M., Gallagher, S.M., Warren, S.T., and Bear, M.F. (2002). Altered synaptic plasticity in a mouse model of fragile X mental retardation. *Proc Natl Acad Sci U S A* 99**,** 7746-7750.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., Gottardo, R., Hahne, F., Hansen, K.D., Irizarry, R.A., Lawrence, M., Love, M.I., Macdonald, J., Obenchain, V., Oles, A.K., Pages, H., Reyes, A., Shannon, P., Smyth, G.K., Tenenbaum, D., Waldron, L., and Morgan, M. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 12**,** 115-121.

Huppertz, I., Attig, J., D'ambrogio, A., Easton, L.E., Sibley, C.R., Sugimoto, Y., Tajnik, M., Konig, J., and Ule, J. (2014). iCLIP: protein-RNA interactions at nucleotide resolution. *Methods* 65**,** 274-287.

Hussain, S., and Bashir, Z.I. (2015). The epitranscriptome in modulating spatiotemporal RNA translation in neuronal post-synaptic function. *Front Cell Neurosci* 9**,** 420.

Hybels, C.F., Blazer, D.G., Landerman, L.R., and Steffens, D.C. (2011). Heterogeneity in symptom profiles among older adults diagnosed with major depression. *Int Psychogeriatr* 23**,** 906-922.

Ibrahim, L., Diazgranados, N., Luckenbaugh, D.A., Machado-Vieira, R., Baumann, J., Mallinger, A.G., and Zarate, C.A., Jr. (2011). Rapid decrease in depressive symptoms with an N-methyl-d-aspartate antagonist in ECT-resistant major depression. *Prog Neuropsychopharmacol Biol Psychiatry* 35**,** 1155-1159.

Ifrim, M.F., Williams, K.R., and Bassell, G.J. (2015). Single-Molecule Imaging of PSD-95 mRNA Translation in Dendrites and Its Dysregulation in a Mouse Model of Fragile X Syndrome. *J Neurosci* 35**,** 7116-7130.

Ince-Dunn, G., Okano, H.J., Jensen, K.B., Park, W.Y., Zhong, R., Ule, J., Mele, A., Fak, J.J., Yang, C., Zhang, C., Yoo, J., Herre, M., Okano, H., Noebels, J.L., and Darnell, R.B. (2012). Neuronal Elav-like (Hu) proteins regulate RNA splicing and abundance to control glutamate levels and neuronal excitability. *Neuron* 75**,** 1067-1080.

Ingolia, N.T., Brar, G.A., Rouskin, S., Mcgeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat Protoc* 7**,** 1534-1550.

Ingolia, N.T., Ghaemmaghami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324**,** 218-223.

Inoki, K., Ouyang, H., Zhu, T., Lindvall, C., Wang, Y., Zhang, X., Yang, Q., Bennett, C., Harada, Y., Stankunas, K., Wang, C.Y., He, X., Macdougald, O.A., You, M., Williams, B.O., and Guan, K.L. (2006). TSC2 integrates Wnt and energy signals via a coordinated phosphorylation by AMPK and GSK3 to regulate cell growth. *Cell* 126**,** 955-968.

Inoki, K., Zhu, T., and Guan, K.L. (2003). TSC2 mediates cellular energy response to control cell growth and survival. *Cell* 115**,** 577-590.

Jain, R., Devine, T., George, A.D., Chittur, S.V., Baroni, T.E., Penalva, L.O., and Tenenbaum, S.A. (2011). RIP-Chip analysis: RNA-Binding Protein Immunoprecipitation-Microarray (Chip) Profiling. *Methods Mol Biol* 703**,** 247-263.

Jaso, B.A., Niciu, M.J., Iadarola, N.D., Lally, N., Richards, E.M., Park, M., Ballard, E.D., Nugent, A.C., Machado-Vieira, R., and Zarate, C.A. (2017). Therapeutic Modulation of Glutamate Receptors in Major Depressive Disorder. *Curr Neuropharmacol* 15**,** 57-70.

Jembrek, M.J., and Vlainic, J. (2015). GABA Receptors: Pharmacological Potential and Pitfalls. *Curr Pharm Des* 21**,** 4943-4959.

Jernigan, C.S., Goswami, D.B., Austin, M.C., Iyo, A.H., Chandran, A., Stockmeier, C.A., and Karolewicz, B. (2011). The mTOR signaling pathway in the prefrontal cortex is compromised in major depressive disorder. *Prog Neuropsychopharmacol Biol Psychiatry* 35**,** 1774-1779.

Jia, J., and Le, W. (2015). Molecular network of neuronal autophagy in the pathophysiology and treatment of depression. *Neurosci Bull* 31**,** 427-434.

Jiao, Y., and Meyerowitz, E.M. (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. *Mol Syst Biol* 6**,** 419.

Jin, J., Li, G.J., Davis, J., Zhu, D., Wang, Y., Pan, C., and Zhang, J. (2007). Identification of novel proteins associated with both alpha-synuclein and DJ-1. *Mol Cell Proteomics* 6**,** 845-859.

Job, C., and Eberwine, J. (2001). Identification of sites for exponential translation in living dendrites. *Proc Natl Acad Sci U S A* 98**,** 13037-13042.

Jobim, P.F., Pedroso, T.R., Christoff, R.R., Werenicz, A., Maurmann, N., Reolon, G.K., and Roesler, R. (2012). Inhibition of mTOR by rapamycin in the amygdala or hippocampus impairs formation and reconsolidation of inhibitory avoidance memory. *Neurobiol Learn Mem* 97**,** 105-112.

Ju, W., Morishita, W., Tsui, J., Gaietta, G., Deerinck, T.J., Adams, S.R., Garner, C.C., Tsien, R.Y., Ellisman, M.H., and Malenka, R.C. (2004). Activity-dependent

regulation of dendritic synthesis and trafficking of AMPA receptors. *Nat Neurosci* 7**,** 244-253.

Jung, C.H., Jun, C.B., Ro, S.H., Kim, Y.M., Otto, N.M., Cao, J., Kundu, M., and Kim, D.H. (2009). ULK-Atg13-FIP200 complexes mediate mTOR signaling to the autophagy machinery. *Mol Biol Cell* 20**,** 1992-2003.

Jung, H., and Holt, C.E. (2011). Local translation of mRNAs in neural development. *Wiley Interdiscip Rev RNA* 2**,** 153-165.

Jungkamp, A.C., Stoeckius, M., Mecenas, D., Grun, D., Mastrobuoni, G., Kempa, S., and Rajewsky, N. (2011). In vivo and transcriptome-wide identification of RNA binding protein target sites. *Mol Cell* 44**,** 828-840.

Kaeberlein, M. (2013). mTOR Inhibition: From Aging to Autism and Beyond. *Scientifica (Cairo)* 2013**,** 849186.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* 42**,** D199-205.

Kang, H., and Schuman, E.M. (1996). A requirement for local protein synthesis in neurotrophin-induced hippocampal synaptic plasticity. *Science* 273**,** 1402-1406.

Kang, H.J., Voleti, B., Hajszan, T., Rajkowska, G., Stockmeier, C.A., Licznerski, P., Lepack, A., Majik, M.S., Jeong, L.S., Banasr, M., Son, H., and Duman, R.S. (2012). Decreased expression of synapse-related genes and loss of synapses in major depressive disorder. *Nat Med* 18**,** 1413-1417.

Kassambara, A., and Mundt, F. (2016). "factoextra: Extract and Visualize the Results of Multivariate Data Analyses.". R package version 1.0.3 ed.).

Katche, C., Cammarota, M., and Medina, J.H. (2013). Molecular signatures and mechanisms of long-lasting memory consolidation and storage. *Neurobiol Learn Mem* 106**,** 40-47.

Katz, Z.B., English, B.P., Lionnet, T., Yoon, Y.J., Monnier, N., Ovryn, B., Bathe, M., and Singer, R.H. (2016). Mapping translation 'hot-spots' in live cells by tracking single molecules of mRNA and ribosomes. *Elife* 5.

Kavalali, E.T., and Monteggia, L.M. (2012). Synaptic mechanisms underlying rapid antidepressant action of ketamine. *Am J Psychiatry* 169**,** 1150-1156.

Kavalali, E.T., and Monteggia, L.M. (2015). How does ketamine elicit a rapid antidepressant response? *Curr Opin Pharmacol* 20**,** 35-39.

Keck, T., Keller, G.B., Jacobsen, R.I., Eysel, U.T., Bonhoeffer, T., and Hubener, M. (2013). Synaptic scaling and homeostatic plasticity in the mouse visual cortex in vivo. *Neuron* 80**,** 327-334.

Keck, T., Scheuss, V., Jacobsen, R.I., Wierenga, C.J., Eysel, U.T., Bonhoeffer, T., and Hubener, M. (2011). Loss of sensory input causes rapid structural changes of inhibitory neurons in adult mouse visual cortex. *Neuron* 71**,** 869-882.

Keene, J.D., Komisarow, J.M., and Friedersdorf, M.B. (2006). RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat Protoc* 1**,** 302-307.

Kegel, L., Aunin, E., Meijer, D., and Bermingham, J.R. (2013). LGI proteins in the nervous system. *ASN Neuro* 5**,** 167-181.

Kelleher, R.J., 3rd, and Bear, M.F. (2008). The autistic neuron: troubled translation? *Cell* 135**,** 401-406.

Kelleher, R.J., 3rd, Govindarajan, A., Jung, H.Y., Kang, H., and Tonegawa, S. (2004). Translational control by MAPK signaling in long-term synaptic plasticity and memory. *Cell* 116**,** 467-479.

Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal Chem* 74**,** 5383-5392.

Kenney, J.W., Genheden, M., Moon, K.M., Wang, X., Foster, L.J., and Proud, C.G. (2016). Eukaryotic elongation factor 2 kinase regulates the synthesis of microtubule-related proteins in neurons. *J Neurochem* 136**,** 276-284.

Kenny, P.J. (2014). Epigenetics, microRNA, and addiction. *Dialogues Clin Neurosci* 16**,** 335-344.

Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R., Rush, A.J., Walters, E.E., Wang, P.S., and National Comorbidity Survey, R. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA* 289**,** 3095-3105.

Kiebler, M.A., and Bassell, G.J. (2006). Neuronal RNA granules: movers and makers. *Neuron* 51**,** 685-690.

Kim, C.S., Chang, P.Y., and Johnston, D. (2012). Enhancement of dorsal hippocampal activity by knockdown of HCN1 channels leads to anxiolytic- and antidepressant-like behaviors. *Neuron* 75**,** 503-516.

Kim, D., Langmead, B., and Salzberg, S.L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat Methods* 12**,** 357-360.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013a). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* 14**,** R36.

Kim, D.H., Sarbassov, D.D., Ali, S.M., King, J.E., Latek, R.R., Erdjument-Bromage, H., Tempst, P., and Sabatini, D.M. (2002). mTOR interacts with raptor to form a nutrient-sensitive complex that signals to the cell growth machinery. *Cell* 110**,** 163-175.

Kim, J., and Hoffman, D.A. (2008). Potassium channels: newly found players in synaptic plasticity. *Neuroscientist* 14**,** 276-286.

Kim, K.K., Yang, Y., Zhu, J., Adelstein, R.S., and Kawamoto, S. (2014). Rbfox3 controls the biogenesis of a subset of microRNAs. *Nat Struct Mol Biol* 21**,** 901-910.

Kim, T.K., Sul, J.Y., Helmfors, H., Langel, U., Kim, J., and Eberwine, J. (2013b). Dendritic glutamate receptor mRNAs show contingent local hotspot-dependent translational dynamics. *Cell Rep* 5**,** 114-125.

Kirkwood, A., Rioult, M.C., and Bear, M.F. (1996). Experience-dependent modification of synaptic plasticity in visual cortex. *Nature* 381**,** 526-528.

Kirkwood, K.J., Ahmad, Y., Larance, M., and Lamond, A.I. (2013). Characterization of native protein complexes and protein isoform variation using size-fractionation-based quantitative proteomics. *Mol Cell Proteomics* 12**,** 3851-3873.

Kirov, S.A., and Harris, K.M. (1999). Dendrites are more spiny on mature hippocampal neurons when synapses are inactivated. *Nat Neurosci* 2**,** 878-883.

Kishore, S., Jaskiewicz, L., Burger, L., Hausser, J., Khorshid, M., and Zavolan, M. (2011). A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat Methods* 8**,** 559-564.

Kitchen, R.R., Rozowsky, J.S., Gerstein, M.B., and Nairn, A.C. (2014). Decoding neuroproteomics: integrating the genome, translatome and functional anatomy. *Nat Neurosci* 17**,** 1491-1499.

Klinke, D.J., 2nd (2009). An empirical Bayesian approach for model-based inference of cellular signaling networks. *BMC Bioinformatics* 10**,** 371.

Kolde, R. (2015). "pheatmap: Pretty Heatmaps.". R package version 1.0.8 ed.).

Konig, J., Zarnack, K., Rot, G., Curk, T., Kayikci, M., Zupan, B., Turner, D.J., Luscombe, N.M., and Ule, J. (2010). iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* 17**,** 909-915.

Koob, G.F., and Le Moal, M. (2001). Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* 24**,** 97-129.

Korff, A., Liu, C., Ginghina, C., Shi, M., Zhang, J., and Alzheimer's Disease Neuroimaging, I. (2013). alpha-Synuclein in cerebrospinal fluid of Alzheimer's disease and mild cognitive impairment. *J Alzheimers Dis* 36**,** 679-688.

Kos, A., Wanke, K.A., Gioio, A., Martens, G.J., Kaplan, B.B., and Aschrafi, A. (2016). Monitoring mRNA Translation in Neuronal Processes Using Fluorescent Non-Canonical Amino Acid Tagging. *J Histochem Cytochem* 64**,** 323-333.

Kosten, T.A., Galloway, M.P., Duman, R.S., Russell, D.S., and D'sa, C. (2008). Repeated unpredictable stress and antidepressants differentially regulate expression of the bcl-2 family of apoptotic genes in rat cortical, hippocampal, and limbic brain structures. *Neuropsychopharmacology* 33**,** 1545-1558.

Krishnan-Sarin, S., O'malley, S., and Krystal, J.H. (2008). Treatment implications: using neuroscience to guide the development of new pharmacotherapies for alcoholism. *Alcohol Res Health* 31**,** 400-407.

Krishnan, V., and Nestler, E.J. (2008). The molecular neurobiology of depression. *Nature* 455**,** 894-902.

Kudla, G., Granneman, S., Hahn, D., Beggs, J.D., and Tollervey, D. (2011). Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc Natl Acad Sci U S A* 108**,** 10010-10015.

Kulkarni, M.M. (2011). Digital multiplexed gene expression analysis using the NanoString nCounter system. *Curr Protoc Mol Biol* Chapter 25**,** Unit25B 10.

Kundu, P., Fabian, M.R., Sonenberg, N., Bhattacharyya, S.N., and Filipowicz, W. (2012). HuR protein attenuates miRNA-mediated repression by promoting miRISC dissociation from the target RNA. *Nucleic Acids Res* 40**,** 5088-5100.

Kye, M.J., Liu, T., Levy, S.F., Xu, N.L., Groves, B.B., Bonneau, R., Lao, K., and Kosik, K.S. (2007). Somatodendritic microRNAs identified by laser capture and multiplex RT-PCR. *RNA* 13**,** 1224-1234.

Lagier-Tourenne, C., Polymenidou, M., Hutt, K.R., Vu, A.Q., Baughn, M., Huelga, S.C., Clutario, K.M., Ling, S.C., Liang, T.Y., Mazur, C., Wancewicz, E., Kim, A.S., Watt, A., Freier, S., Hicks, G.G., Donohue, J.P., Shiue, L., Bennett, C.F., Ravits, J., Cleveland, D.W., and Yeo, G.W. (2012). Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat Neurosci* 15**,** 1488-1497.

Lahens, N.F., Kavakli, I.H., Zhang, R., Hayer, K., Black, M.B., Dueck, H., Pizarro, A., Kim, J., Irizarry, R., Thomas, R.S., Grant, G.R., and Hogenesch, J.B. (2014). IVT-seq reveals extreme bias in RNA sequencing. *Genome Biol* 15**,** R86.

Lamb, J. (2007). The Connectivity Map: a new tool for biomedical research. *Nat Rev Cancer* 7**,** 54-60.

Lamb, J., Crawford, E.D., Peck, D., Modell, J.W., Blat, I.C., Wrobel, M.J., Lerner, J., Brunet, J.P., Subramanian, A., Ross, K.N., Reich, M., Hieronymus, H., Wei, G., Armstrong, S.A., Haggarty, S.J., Clemons, P.A., Wei, R., Carr, S.A., Lander, E.S., and Golub, T.R. (2006). The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313**,** 1929-1935.

Lambert, N., Robertson, A., Jangi, M., Mcgeary, S., Sharp, P.A., and Burge, C.B. (2014). RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. *Mol Cell* 54**,** 887-900.

Lan, A.P., Chen, J., Zhao, Y., Chai, Z., and Hu, Y. (2016). mTOR Signaling in Parkinson's Disease. *Neuromolecular Med*.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9**,** 357-359.

Laplante, M., and Sabatini, D.M. (2009). mTOR signaling at a glance. *J Cell Sci* 122**,** 3589-3594.

Laplante, M., and Sabatini, D.M. (2012). mTOR signaling in growth control and disease. *Cell* 149**,** 274-293.

Lau, C.G., Takayasu, Y., Rodenas-Ruano, A., Paternain, A.V., Lerma, J., Bennett, M.V., and Zukin, R.S. (2010). SNAP-25 is a target of protein kinase C phosphorylation critical to NMDA receptor trafficking. *J Neurosci* 30**,** 242-254.

Lavallee-Adam, M., Park, S.K., Martinez-Bartolome, S., He, L., and Yates, J.R., 3rd (2015). From raw data to biological discoveries: a computational analysis pipeline for mass spectrometry-based proteomics. *J Am Soc Mass Spectrom* 26**,** 1820-1826.

Law, C.W., Chen, Y., Shi, W., and Smyth, G.K. (2014). voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* 15**,** R29.

Lawrence, M., Huber, W., Pages, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M.T., and Carey, V.J. (2013). Software for computing and annotating genomic ranges. *PLoS Comput Biol* 9**,** e1003118.

Lecuyer, E., Yoshida, H., Parthasarathy, N., Alm, C., Babak, T., Cerovina, T., Hughes, T.R., Tomancak, P., and Krause, H.M. (2007). Global analysis of mRNA

localization reveals a prominent role in organizing cellular architecture and function. *Cell* 131**,** 174-187.

Lee, B.H., Smith, T., and Paciorkowski, A.R. (2015). Autism spectrum disorder and epilepsy: Disorders with a shared biology. *Epilepsy Behav* 47**,** 191-201.

Lee, C.C., Huang, C.C., Wu, M.Y., and Hsu, K.S. (2005). Insulin stimulates postsynaptic density-95 protein translation via the phosphoinositide 3-kinase-Akt-mammalian target of rapamycin signaling pathway. *J Biol Chem* 280**,** 18543-18550.

Lee, D.F., Kuo, H.P., Chen, C.T., Hsu, J.M., Chou, C.K., Wei, Y., Sun, H.L., Li, L.Y., Ping, B., Huang, W.C., He, X., Hung, J.Y., Lai, C.C., Ding, Q., Su, J.L., Yang, J.Y., Sahin, A.A., Hortobagyi, G.N., Tsai, F.J., Tsai, C.H., and Hung, M.C. (2007). IKK beta suppression of TSC1 links inflammation and tumor angiogenesis via the mTOR pathway. *Cell* 130**,** 440-455.

Lee, E.J., and Tournier, C. (2011). The requirement of uncoordinated 51-like kinase 1 (ULK1) and ULK2 in the regulation of autophagy. *Autophagy* 7**,** 689-695.

Lee, H.Y., Ge, W.P., Huang, W., He, Y., Wang, G.X., Rowson-Baldwin, A., Smith, S.J., Jan, Y.N., and Jan, L.Y. (2011). Bidirectional regulation of dendritic voltage-gated potassium channels by the fragile X mental retardation protein. *Neuron* 72**,** 630-642.

Lee, M.C., Yasuda, R., and Ehlers, M.D. (2010a). Metaplasticity at single glutamatergic synapses. *Neuron* 66**,** 859-870.

Lee, S.Y., Hahn, C.Y., Lee, J.F., Huang, S.Y., Chen, S.L., Kuo, P.H., Lee, I.H., Yeh, T.L., Yang, Y.K., Chen, S.H., Ko, H.C., and Lu, R.B. (2010b). MAOA interacts with the ALDH2 gene in anxiety-depression alcohol dependence. *Alcohol Clin Exp Res* 34**,** 1212-1218.

Leng, N., Dawson, J.A., Thomson, J.A., Ruotti, V., Rissman, A.I., Smits, B.M., Haag, J.D., Gould, M.N., Stewart, R.M., and Kendziorski, C. (2013). EBSeq: an empirical Bayes hierarchical model for inference in RNA-seq experiments. *Bioinformatics* 29**,** 1035-1043.

Leng, N., Li, Y., Mcintosh, B.E., Nguyen, B.K., Duffin, B., Tian, S., Thomson, J.A., Dewey, C.N., Stewart, R., and Kendziorski, C. (2015). EBSeq-HMM: a Bayesian approach for identifying gene-expression changes in ordered RNA-seq experiments. *Bioinformatics* 31**,** 2614-2622.

Lenzken, S.C., Achsel, T., Carri, M.T., and Barabino, S.M. (2014). Neuronal RNA-binding proteins in health and disease. *Wiley Interdiscip Rev RNA* 5**,** 565-576.

Lepack, A.E., Fuchikami, M., Dwyer, J.M., Banasr, M., and Duman, R.S. (2014). BDNF release is required for the behavioral actions of ketamine. *Int J Neuropsychopharmacol* 18.

Leung, K.M., and Holt, C.E. (2008). Live visualization of protein synthesis in axonal growth cones by microinjection of photoconvertible Kaede into Xenopus embryos. *Nat Protoc* 3**,** 1318-1327.

Leung, K.M., Van Horck, F.P., Lin, A.C., Allison, R., Standart, N., and Holt, C.E. (2006). Asymmetrical beta-actin mRNA translation in growth cones mediates attractive turning to netrin-1. *Nat Neurosci* 9**,** 1247-1256.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009a). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25**,** 2078-2079.

Li, N., Lee, B., Liu, R.J., Banasr, M., Dwyer, J.M., Iwata, M., Li, X.Y., Aghajanian, G., and Duman, R.S. (2010). mTOR-dependent synapse formation underlies the rapid antidepressant effects of NMDA antagonists. *Science* 329**,** 959-964.

Li, N., Liu, R.J., Dwyer, J.M., Banasr, M., Lee, B., Son, H., Li, X.Y., Aghajanian, G., and Duman, R.S. (2011). Glutamate N-methyl-D-aspartate receptor antagonists rapidly reverse behavioral and synaptic deficits caused by chronic stress exposure. *Biol Psychiatry* 69**,** 754-761.

Li, R., Dong, Q., Yuan, X., Zeng, X., Gao, Y., Chiao, C., Li, H., Zhao, X., Keles, S., Wang, Z., and Chang, Q. (2016). Misregulation of Alternative Splicing in a Mouse Model of Rett Syndrome. *PLoS Genet* 12**,** e1006129.

Li, R., Yu, C., Li, Y., Lam, T.W., Yiu, S.M., Kristiansen, K., and Wang, J. (2009b). SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics* 25**,** 1966-1967.

Li, X., Alafuzoff, I., Soininen, H., Winblad, B., and Pei, J.J. (2005). Levels of mTOR and its downstream targets 4E-BP1, eEF2, and eEF2 kinase in relationships with tau in Alzheimer's disease brain. *FEBS J* 272**,** 4211-4220.

Li, X., Nair, A., Wang, S., and Wang, L. (2015). Quality control of RNA-seq experiments. *Methods Mol Biol* 1269**,** 137-146.

Li, X., Zhao, X., Fang, Y., Jiang, X., Duong, T., Fan, C., Huang, C.C., and Kain, S.R. (1998). Generation of destabilized green fluorescent protein as a transcription reporter. *J Biol Chem* 273**,** 34970-34975.

Liao, Y., Smyth, G.K., and Shi, W. (2013). The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res* 41**,** e108.

Licatalosi, D.D., Mele, A., Fak, J.J., Ule, J., Kayikci, M., Chi, S.W., Clark, T.A., Schweitzer, A.C., Blume, J.E., Wang, X., Darnell, J.C., and Darnell, R.B. (2008). HITS-CLIP yields genome-wide insights into brain alternative RNA processing. *Nature* 456**,** 464-469.

Lieblich, S.M., Castle, D.J., Pantelis, C., Hopwood, M., Young, A.H., and Everall, I.P. (2015). High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. *BJPsych Open* 1**,** e5-e7.

Lin, M.Z., Glenn, J.S., and Tsien, R.Y. (2008). A drug-controllable tag for visualizing newly synthesized proteins in cells and whole animals. *Proc Natl Acad Sci U S A* 105**,** 7744-7749.

Lin, M.Z., and Tsien, R.Y. (2010). TimeSTAMP tagging of newly synthesized proteins. *Curr Protoc Protein Sci* Chapter 26**,** Unit 26 25.

Lin, Y., Golovnina, K., Chen, Z.X., Lee, H.N., Negron, Y.L., Sultana, H., Oliver, B., and Harbison, S.T. (2016). Comparison of normalization and differential expression analyses using RNA-Seq data from 726 individual Drosophila melanogaster. *BMC Genomics* 17**,** 28.

Lipton, J.O., and Sahin, M. (2014). The neurology of mTOR. *Neuron* 84**,** 275-291.

Liu, B., Archer, C.T., Burdine, L., Gillette, T.G., and Kodadek, T. (2007). Label transfer chemistry for the characterization of protein-protein interactions. *J Am Chem Soc* 129**,** 12348-12349.

Liu, H.H., and Cline, H.T. (2016). Fragile X Mental Retardation Protein Is Required to Maintain Visual Conditioning-Induced Behavioral Plasticity by Limiting Local Protein Synthesis. *J Neurosci* 36**,** 7325-7339.

Liu, R.J., Lee, F.S., Li, X.Y., Bambico, F., Duman, R.S., and Aghajanian, G.K. (2012). Brain-derived neurotrophic factor Val66Met allele impairs basal and ketamine-stimulated synaptogenesis in prefrontal cortex. *Biol Psychiatry* 71**,** 996-1005.

Liu, Y., Ferguson, J.F., Xue, C., Silverman, I.M., Gregory, B., Reilly, M.P., and Li, M. (2013). Evaluating the impact of sequencing depth on transcriptome profiling in human adipose. *PLoS One* 8**,** e66883.

Liu, Y., Zhou, J., and White, K.P. (2014). RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics* 30**,** 301-304.

Liu, Z., Huang, L., Luo, X.J., Wu, L., and Li, M. (2016). MAOA Variants and Genetic Susceptibility to Major Psychiatric Disorders. *Mol Neurobiol* 53**,** 4319-4327.

Livingstone, M., Atas, E., Meller, A., and Sonenberg, N. (2010). Mechanisms governing the control of mRNA translation. *Phys Biol* 7**,** 021001.

Lohoff, F.W. (2010). Overview of the genetics of major depressive disorder. *Curr Psychiatry Rep* 12**,** 539-546.

Loman, N.J., Misra, R.V., Dallman, T.J., Constantinidou, C., Gharbia, S.E., Wain, J., and Pallen, M.J. (2012). Performance comparison of benchtop high-throughput sequencing platforms. *Nat Biotechnol* 30**,** 434-439.

Long, X., Lin, Y., Ortiz-Vega, S., Yonezawa, K., and Avruch, J. (2005). Rheb binds and regulates the mTOR kinase. *Curr Biol* 15**,** 702-713.

Loos, M., Li, K.W., Van Der Schors, R., Gouwenberg, Y., Van Der Loo, R., Williams, R.W., Smit, A.B., and Spijker, S. (2016). Impact of genetic variation on synaptic protein levels in genetically diverse mice. *Proteomics* 16**,** 1123-1130.

Lopez, A.D., Mathers, C.D., Ezzati, M., Jamison, D.T., and Murray, C.J.L. (2006). "Measuring the Global Burden of Disease and Risk Factors, 1990-2001," in *Global Burden of Disease and Risk Factors,* eds. A.D. Lopez, C.D. Mathers, M. Ezzati, D.T. Jamison & C.J.L. Murray. (Washington (DC)).

Lovatt, D., Bell, T., and Eberwine, J. (2015). Single-neuron isolation for RNA analysis using pipette capture and laser capture microdissection. *Cold Spring Harb Protoc* 2015**,** pdb prot072439.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15**,** 550.

Lovinger, D.M., and Roberto, M. (2013). Synaptic effects induced by alcohol. *Curr Top Behav Neurosci* 13**,** 31-86.

Lovinger, D.M., White, G., and Weight, F.F. (1989). Ethanol inhibits NMDA-activated ion current in hippocampal neurons. *Science* 243**,** 1721-1724.

Lovinger, D.M., White, G., and Weight, F.F. (1990). NMDA receptor-mediated synaptic excitation selectively inhibited by ethanol in hippocampal slice from adult rat. *J Neurosci* 10**,** 1372-1379.

Lu, Y., Ferris, J., and Gao, F.B. (2009). Frontotemporal dementia and amyotrophic lateral sclerosis-associated disease protein TDP-43 promotes dendritic branching. *Mol Brain* 2**,** 30.

Lu, Z., Guan, X., Schmidt, C.A., and Matera, A.G. (2014). RIP-seq analysis of eukaryotic Sm proteins identifies three major categories of Sm-containing ribonucleoproteins. *Genome Biol* 15**,** R7.

Lucas, A. (2014). "amap: Another Multidimensional Analysis Package". R package version 0.8-14 ed.).

Luchelli, L., Thomas, M.G., and Boccaccio, G.L. (2015). Synaptic control of mRNA translation by reversible assembly of XRN1 bodies. *J Cell Sci* 128**,** 1542-1554.

Luikart, B.W., Schnell, E., Washburn, E.K., Bensen, A.L., Tovar, K.R., and Westbrook, G.L. (2011). Pten knockdown in vivo increases excitatory drive onto dentate granule cells. *J Neurosci* 31**,** 4345-4354.

Lukyanov, K.A., Chudakov, D.M., Lukyanov, S., and Verkhusha, V.V. (2005). Innovation: Photoactivatable fluorescent proteins. *Nat Rev Mol Cell Biol* 6**,** 885-891.

Luo, K.R., Hong, C.J., Liou, Y.J., Hou, S.J., Huang, Y.H., and Tsai, S.J. (2010). Differential regulation of neurotrophin S100B and BDNF in two rat models of depression. *Prog Neuropsychopharmacol Biol Psychiatry* 34**,** 1433-1439.

Luscher, C., and Malenka, R.C. (2011). Drug-evoked synaptic plasticity in addiction: from molecular changes to circuit remodeling. *Neuron* 69**,** 650-663.

Lux, V., and Kendler, K.S. (2010). Deconstructing major depression: a validation study of the DSM-IV symptomatic criteria. *Psychol Med* 40**,** 1679-1690.

Ma, T., Trinh, M.A., Wexler, A.J., Bourbon, C., Gatti, E., Pierre, P., Cavener, D.R., and Klann, E. (2013). Suppression of eIF2alpha kinases alleviates Alzheimer's disease-related plasticity and memory deficits. *Nat Neurosci* 16**,** 1299-1305.

Maier, T., Guell, M., and Serrano, L. (2009). Correlation of mRNA and protein in complex biological samples. *FEBS Lett* 583**,** 3966-3973.

Manger, P.R., Cort, J., Ebrahim, N., Goodman, A., Henning, J., Karolia, M., Rodrigues, S.L., and Strkalj, G. (2008). Is 21st century neuroscience too focussed on the rat/mouse model of brain function and dysfunction? *Front Neuroanat* 2**,** 5.

Manninen, T., Hituri, K., Kotaleski, J.H., Blackwell, K.T., and Linne, M.L. (2010). Postsynaptic signal transduction models for long-term potentiation and depression. *Front Comput Neurosci* 4**,** 152.

Marcel, M. (2011). Cutadapt removes sequences from high-throughput sequencing reads. *EMBnet.journal* 17**,** 3.

Margeta-Mitrovic, M., Jan, Y.N., and Jan, L.Y. (2000). A trafficking checkpoint controls GABA(B) receptor heterodimerization. *Neuron* 27**,** 97-106.

Marino, G., Niso-Santano, M., Baehrecke, E.H., and Kroemer, G. (2014). Self-consumption: the interplay of autophagy and apoptosis. *Nat Rev Mol Cell Biol* 15**,** 81-94.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., and Gilad, Y. (2008). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18**,** 1509-1517.

Markham, K., Bai, Y., and Schmitt-Ulms, G. (2007). Co-immunoprecipitations revisited: an update on experimental concepts and their implementation for sensitive interactome investigations of endogenous proteins. *Anal Bioanal Chem* 389**,** 461-473.

Martin, A., Ochagavia, M.E., Rabasa, L.C., Miranda, J., Fernandez-De-Cossio, J., and Bringas, R. (2010). BisoGenet: a new tool for gene network building, visualization and analysis. *BMC Bioinformatics* 11**,** 91.

Martin, B.R., Giepmans, B.N., Adams, S.R., and Tsien, R.Y. (2005). Mammalian cell-based optimization of the biarsenical-binding tetracysteine motif for improved fluorescence and affinity. *Nat Biotechnol* 23**,** 1308-1314.

Martin, K.C., Casadio, A., Zhu, H., Yaping, E., Rose, J.C., Chen, M., Bailey, C.H., and Kandel, E.R. (1997). Synapse-specific, long-term facilitation of aplysia sensory to motor synapses: a function for local protein synthesis in memory storage. *Cell* 91**,** 927-938.

Martin, K.C., and Ephrussi, A. (2009). mRNA localization: gene expression in the spatial dimension. *Cell* 136**,** 719-730.

Martin, K.C., and Zukin, R.S. (2006). RNA trafficking and local protein synthesis in dendrites: an overview. *J Neurosci* 26**,** 7131-7134.

Martin, S.J., Grimwood, P.D., and Morris, R.G. (2000). Synaptic plasticity and memory: an evaluation of the hypothesis. *Annu Rev Neurosci* 23**,** 649-711.

Masek, T., Valasek, L., and Pospisek, M. (2011). Polysome analysis and RNA purification from sucrose gradients. *Methods Mol Biol* 703**,** 293-309.

Mathers, C.D., Lopez, A.D., and Murray, C.J.L. (2006). "The Burden of Disease and Mortality by Condition: Data, Methods, and Results for 2001," in *Global Burden of Disease and Risk Factors,* eds. A.D. Lopez, C.D. Mathers, M. Ezzati, D.T. Jamison & C.J.L. Murray. (Washington (DC)).

Mcclung, C.A. (2007). Circadian genes, rhythms and the biology of mood disorders. *Pharmacol Ther* 114**,** 222-232.

Mcclung, C.A., and Nestler, E.J. (2008). Neuroplasticity mediated by altered gene expression. *Neuropsychopharmacology* 33**,** 3-17.

Mckernan, D.P., Dinan, T.G., and Cryan, J.F. (2009). "Killing the Blues": a role for cellular suicide (apoptosis) in depression and the antidepressant response? *Prog Neurobiol* 88**,** 246-263.

Mcnamara, R.K., and Routtenberg, A. (1995). NMDA receptor blockade prevents kainate induction of protein F1/GAP-43 mRNA in hippocampal granule cells and subsequent mossy fiber sprouting in the rat. *Brain Res Mol Brain Res* 33**,** 22-28.

Megias, M., Emri, Z., Freund, T.F., and Gulyas, A.I. (2001). Total number and distribution of inhibitory and excitatory synapses on hippocampal CA1 pyramidal cells. *Neuroscience* 102**,** 527-540.

Merico, D., Isserlin, R., and Bader, G.D. (2011). Visualizing gene-set enrichment results using the Cytoscape plug-in enrichment map. *Methods Mol Biol* 781**,** 257-277.

Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5**,** e13984.

Michmizos, D., Koutsouraki, E., Asprodini, E., and Baloyannis, S. (2011). Synaptic plasticity: a unifying model to address some persisting questions. *Int J Neurosci* 121**,** 289-304.

Mikula, M., Rubel, T., Karczmarski, J., Statkiewicz, M., Bomsztyk, K., and Ostrowski, J. (2015). Beads-free protein immunoprecipitation for a mass spectrometry-based interactome and posttranslational modifications analysis. *Proteome Sci* 13**,** 23.

Mili, S., and Steitz, J.A. (2004). Evidence for reassociation of RNA-binding proteins after cell lysis: implications for the interpretation of immunoprecipitation analyses. *RNA* 10**,** 1692-1694.

Miller-Fleming, L., Antas, P., Pais, T.F., Smalley, J.L., Giorgini, F., and Outeiro, T.F. (2014). Yeast DJ-1 superfamily members are required for diauxic-shift reprogramming and cell survival in stationary phase. *Proc Natl Acad Sci U S A* 111**,** 7012-7017.

Miller, O.H., Yang, L., Wang, C.C., Hargroder, E.A., Zhang, Y., Delpire, E., and Hall, B.J. (2014). GluN2B-containing NMDA receptors regulate depression-like behavior and are critical for the rapid antidepressant actions of ketamine. *Elife* 3**,** e03581.

Miller, S., Yasuda, M., Coats, J.K., Jones, Y., Martone, M.E., and Mayford, M. (2002). Disruption of dendritic translation of CaMKIIalpha impairs stabilization of synaptic plasticity and memory consolidation. *Neuron* 36**,** 507-519.

Miyashiro, K., Dichter, M., and Eberwine, J. (1994). On the nature and differential distribution of mRNAs in hippocampal neurites: implications for neuronal functioning. *Proc Natl Acad Sci U S A* 91**,** 10800-10804.

Miyashiro, K.Y., Beckel-Mitchener, A., Purk, T.P., Becker, K.G., Barret, T., Liu, L., Carbonetto, S., Weiler, I.J., Greenough, W.T., and Eberwine, J. (2003). RNA cargoes associating with FMRP reveal deficits in cellular functioning in Fmr1 null mice. *Neuron* 37**,** 417-431.

Morgan, M., Anders, S., Lawrence, M., Aboyoun, P., Pages, H., and Gentleman, R. (2009). ShortRead: a bioconductor package for input, quality assessment and exploration of high-throughput sequence data. *Bioinformatics* 25**,** 2607-2608.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., and Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5**,** 621-628.

Muddashetty, R.S., Kelic, S., Gross, C., Xu, M., and Bassell, G.J. (2007). Dysregulated metabotropic glutamate receptor-dependent translation of AMPA receptor and postsynaptic density-95 mRNAs at synapses in a mouse model of fragile X syndrome. *J Neurosci* 27**,** 5338-5348.

Muddashetty, R.S., Nalavadi, V.C., Gross, C., Yao, X., Xing, L., Laur, O., Warren, S.T., and Bassell, G.J. (2011). Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling. *Mol Cell* 42**,** 673-688.

Mukhopadhyay, S., Das, D.N., Panda, P.K., Sinha, N., Naik, P.P., Bissoyi, A., Pramanik, K., and Bhutia, S.K. (2015). Autophagy protein Ulk1 promotes mitochondrial apoptosis through reactive oxygen species. *Free Radic Biol Med* 89**,** 311-321.

Murrough, J.W. (2012). Ketamine as a novel antidepressant: from synapse to behavior. *Clin Pharmacol Ther* 91**,** 303-309.

Murrough, J.W., and Charney, D.S. (2012). Is there anything really novel on the antidepressant horizon? *Curr Psychiatry Rep* 14**,** 643-649.

Na, Y., Park, S., Lee, C., Kim, D.K., Park, J.M., Sockanathan, S., Huganir, R.L., and Worley, P.F. (2016). Real-Time Imaging Reveals Properties of Glutamate-Induced Arc/Arg 3.1 Translation in Neuronal Dendrites. *Neuron* 91**,** 561-573.

Nalavadi, V.C., Muddashetty, R.S., Gross, C., and Bassell, G.J. (2012). Dephosphorylation-induced ubiquitination and degradation of FMRP in dendrites: a role in immediate early mGluR-stimulated translation. *J Neurosci* 32**,** 2582-2587.

Namjoshi, S., and Raab-Graham, K.F. (2017). Screening the Molecular Framework Underlying Local Dendritic mRNA Translation. *Front Mol Neurosci* 10.

Napoli, I., Mercaldo, V., Boyl, P.P., Eleuteri, B., Zalfa, F., De Rubeis, S., Di Marino, D., Mohr, E., Massimi, M., Falconi, M., Witke, W., Costa-Mattioli, M., Sonenberg, N., Achsel, T., and Bagni, C. (2008). The fragile X syndrome protein represses activity-dependent translation through CYFIP1, a new 4E-BP. *Cell* 134**,** 1042-1054.

Navarro, D. (2015). "Learning statistics with R: A tutorial for psychology students and other beginners". 0.5 ed.).

Nesvizhskii, A.I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal Chem* 75**,** 4646-4658.

Neuwirth, E. (2014). "RColorBrewer: ColorBrewer Palettes". R package version 1.1-2 ed.).

Niere, F., Namjoshi, S., Song, E., Dilly, G.A., Schoenhard, G., Zemelman, B.V., Mechref, Y., and Raab-Graham, K.F. (2016). Analysis of Proteins That Rapidly Change Upon Mechanistic/Mammalian Target of Rapamycin Complex 1 (mTORC1) Repression Identifies Parkinson Protein 7 (PARK7) as a Novel Protein Aberrantly Expressed in Tuberous Sclerosis Complex (TSC). *Mol Cell Proteomics* 15**,** 426-444.

Niere, F., Wilkerson, J.R., and Huber, K.M. (2012). Evidence for a fragile X mental retardation protein-mediated translational switch in metabotropic glutamate receptor-triggered Arc translation and long-term depression. *J Neurosci* 32**,** 5924-5936.

Nikiforova, V.J., and Willmitzer, L. (2007). Network visualization and network analysis. *EXS* 97**,** 245-275.

Noebels, J. (2011). A perfect storm: Converging paths of epilepsy and Alzheimer's dementia intersect in the hippocampal formation. *Epilepsia* 52 Suppl 1**,** 39-46.

Nosyreva, E., Szabla, K., Autry, A.E., Ryazanov, A.G., Monteggia, L.M., and Kavalali, E.T. (2013). Acute suppression of spontaneous neurotransmission drives synaptic potentiation. *J Neurosci* 33**,** 6990-7002.

Oeffinger, M., Wei, K.E., Rogers, R., Degrasse, J.A., Chait, B.T., Aitchison, J.D., and Rout, M.P. (2007). Comprehensive analysis of diverse ribonucleoprotein complexes. *Nat Methods* 4**,** 951-956.

Ogawa, S., and Kunugi, H. (2015). Inhibitors of Fatty Acid Amide Hydrolase and Monoacylglycerol Lipase: New Targets for Future Antidepressants. *Curr Neuropharmacol* 13**,** 760-775.

Ohgi, Y., Futamura, T., and Hashimoto, K. (2015). Glutamate Signaling in Synaptogenesis and NMDA Receptors as Potential Therapeutic Targets for Psychiatric Disorders. *Curr Mol Med* 15**,** 206-221.

Ohkawa, T., Fukata, Y., Yamasaki, M., Miyazaki, T., Yokoi, N., Takashima, H., Watanabe, M., Watanabe, O., and Fukata, M. (2013). Autoantibodies to epilepsy-related LGI1 in limbic encephalitis neutralize LGI1-ADAM22 interaction and reduce synaptic AMPA receptors. *J Neurosci* 33**,** 18161-18174.

Olivares-Hernandez, R., Bordel, S., and Nielsen, J. (2011). Codon usage variability determines the correlation between proteome and transcriptome fold changes. *BMC Syst Biol* 5**,** 33.

Ong, S.E., Blagoev, B., Kratchmarova, I., Kristensen, D.B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 1**,** 376-386.

Ori, A., Toyama, B.H., Harris, M.S., Bock, T., Iskar, M., Bork, P., Ingolia, N.T., Hetzer, M.W., and Beck, M. (2015). Integrated Transcriptome and Proteome Analyses Reveal Organ-Specific Proteome Deterioration in Old Rats. *Cell Syst* 1**,** 224-237.

Oshlack, A., Robinson, M.D., and Young, M.D. (2010). From RNA-seq reads to differential expression results. *Genome Biol* 11**,** 220.

Oshlack, A., and Wakefield, M.J. (2009). Transcript length bias in RNA-seq data confounds systems biology. *Biol Direct* 4**,** 14.

Osterweil, E.K., Kind, P.C., and Bear, M.F. (2012). Lifting the mood on treating fragile X. *Biol Psychiatry* 72**,** 895-897.

Ostroff, L.E., Botsford, B., Gindina, S., Cowansage, K.K., Ledoux, J.E., Klann, E., and Hoeffer, C. (2017). Accumulation of polyribosomes in dendritic spine heads, but not bases and necks, during memory consolidation depends on cap-dependent translation initiation. *J Neurosci*.

Ostroff, L.E., Cain, C.K., Bedont, J., Monfils, M.H., and Ledoux, J.E. (2010). Fear and safety learning differentially affect synapse size and dendritic translation in the lateral amygdala. *Proc Natl Acad Sci U S A* 107**,** 9418-9423.

Ostroff, L.E., Manzur, M.K., Cain, C.K., and Ledoux, J.E. (2014). Synapses lacking astrocyte appear in the amygdala during consolidation of Pavlovian threat conditioning. *J Comp Neurol* 522**,** 2152-2163.

Ouyang, Y., Rosenstein, A., Kreiman, G., Schuman, E.M., and Kennedy, M.B. (1999). Tetanic stimulation leads to increased accumulation of Ca(2+)/calmodulin-dependent protein kinase II via dendritic protein synthesis in hippocampal neurons. *J Neurosci* 19**,** 7823-7833.

Ovsepian, S.V., and Dolly, J.O. (2011). Dendritic SNAREs add a new twist to the old neuron theory. *Proc Natl Acad Sci U S A* 108**,** 19113-19120.

Padgett, C.L., and Slesinger, P.A. (2010). GABAB receptor coupling to G-proteins and ion channels. *Adv Pharmacol* 58**,** 123-147.

Pages, H., Aboyoun, P., Gentleman, R., and Debroy, S. (2016). "Biostrings: String objects representing biological sequences, and matching algorithms". R package version 2.42.1 ed.).

Pages, H., Carlson, M., Falcon, S., and Li, N. (2017). "AnnotationDbi: Annottion Database Interface". R package version 1.36.2 ed.).

Palacios-Pru, E., Miranda-Contreras, L., Mendoza, R.V., and Zambrano, E. (1988). Dendritic RNA and postsynaptic density formation in chick cerebellar synaptogenesis. *Neuroscience* 24**,** 111-118.

Papoutsi, A., Sidiropoulou, K., and Poirazi, P. (eds.). (2012). *Memory beyond synaptic plasticity: The role of intrinsic neuronal excitability.* Signapore: World Scientific Publishing.

Paradies, M.A., and Steward, O. (1997). Multiple subcellular mRNA distribution patterns in neurons: a nonisotopic in situ hybridization analysis. *J Neurobiol* 33**,** 473-493.

Parsons, R.G., Gafford, G.M., and Helmstetter, F.J. (2006). Translational control via the mammalian target of rapamycin pathway is critical for the formation and stability of long-term fear memory in amygdala neurons. *J Neurosci* 26**,** 12977-12983.

Pe'er, D. (2005). Bayesian network analysis of signaling networks: a primer. *Sci STKE* 2005**,** pl4.

Pei, J.J., and Hugon, J. (2008). mTOR-dependent signalling in Alzheimer's disease. *J Cell Mol Med* 12**,** 2525-2532.

Penalva, L.O., Tenenbaum, S.A., and Keene, J.D. (2004). Gene expression analysis of messenger RNP complexes. *Methods Mol Biol* 257**,** 125-134.

Peters, A., Palay, S.L., and Webster, D.F. (1976). *The Fine Structure of the Nervous System: The Neurons and Supporting Cells.* Philadelphia, Pennsylvania: Saunders.

Peterson, T.R., Laplante, M., Thoreen, C.C., Sancak, Y., Kang, S.A., Kuehl, W.M., Gray, N.S., and Sabatini, D.M. (2009). DEPTOR is an mTOR inhibitor frequently overexpressed in multiple myeloma cells and required for their survival. *Cell* 137**,** 873-886.

Philpot, B.D., Espinosa, J.S., and Bear, M.F. (2003). Evidence for altered NMDA receptor function as a basis for metaplasticity in visual cortex. *J Neurosci* 23**,** 5583-5588.

Pimentel, J., and Boccaccio, G.L. (2014). Translation and silencing in RNA granules: a tale of sand grains. *Front Mol Neurosci* 7**,** 68.

Pittenger, C., and Duman, R.S. (2008). Stress, depression, and neuroplasticity: a convergence of mechanisms. *Neuropsychopharmacology* 33**,** 88-109.

Placzek, A.N., Molfese, D.L., Khatiwada, S., Viana Di Prisco, G., Huang, W., Sidrauski, C., Krnjevic, K., Amos, C.L., Ray, R., Dani, J.A., Walter, P., Salas, R., and Costa-Mattioli, M. (2016). Translational control of nicotine-evoked synaptic potentiation in mice and neuronal responses in human smokers by eIF2alpha. *Elife* 5.

Pocklington, A.J., Cumiskey, M., Armstrong, J.D., and Grant, S.G. (2006). The proteomes of neurotransmitter receptor complexes form modular networks with distributed functionality underlying plasticity and behaviour. *Mol Syst Biol* 2**,** 2006 0023.

Poolos, N.P., and Johnston, D. (2012). Dendritic ion channelopathy in acquired epilepsy. *Epilepsia* 53 Suppl 9**,** 32-40.

Poon, M.M., Choi, S.H., Jamieson, C.A., Geschwind, D.H., and Martin, K.C. (2006). Identification of process-localized mRNAs from cultured rodent hippocampal neurons. *J Neurosci* 26**,** 13390-13399.

Porsolt, R.D., Bertin, A., Blavet, N., Deniel, M., and Jalfre, M. (1979). Immobility induced by forced swimming in rats: effects of agents which modify central catecholamine and serotonin activity. *Eur J Pharmacol* 57**,** 201-210.

Powell, C.M. (2006). Gene targeting of presynaptic proteins in synaptic plasticity and memory: across the great divide. *Neurobiol Learn Mem* 85**,** 2-15.

Pozo, K., and Goda, Y. (2010). Unraveling mechanisms of homeostatic synaptic plasticity. *Neuron* 66**,** 337-351.

Pun, R.Y., Rolle, I.J., Lasarge, C.L., Hosford, B.E., Rosen, J.M., Uhl, J.D., Schmeltzer, S.N., Faulkner, C., Bronson, S.L., Murphy, B.L., Richards, D.A., Holland, K.D., and Danzer, S.C. (2012). Excessive activation of mTOR in postnatally generated granule cells is sufficient to cause epilepsy. *Neuron* 75**,** 1022-1034.

Qin, M., Huang, T., Kader, M., Krych, L., Xia, Z., Burlin, T., Zeidler, Z., Zhao, T., and Smith, C.B. (2015). R-Baclofen Reverses a Social Behavior Deficit and Elevated Protein Synthesis in a Mouse Model of Fragile X Syndrome. *Int J Neuropsychopharmacol* 18.

Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., and Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC Genomics* 13**,** 341.

Queenan, B.N., Lee, K.J., and Pak, D.T. (2012). Wherefore art thou, homeo(stasis)? Functional diversity in homeostatic synaptic plasticity. *Neural Plast* 2012**,** 718203.

Raab-Graham, K.F., Haddick, P.C., Jan, Y.N., and Jan, L.Y. (2006). Activity- and mTOR-dependent suppression of Kv1.1 channel mRNA translation in dendrites. *Science* 314**,** 144-148.

Raab-Graham, K.F., Workman, E.R., Namjoshi, S., and Niere, F. (2016). Pushing the threshold: How NMDAR antagonists induce homeostasis through protein synthesis to remedy depression. *Brain Res* 1647**,** 94-104.

Raman, K. (2010). Construction and analysis of protein-protein interaction networks. *Autom Exp* 2**,** 2.

Rao, A., and Steward, O. (1991a). Evidence that protein constituents of postsynaptic membrane specializations are locally synthesized: analysis of proteins synthesized within synaptosomes. *J Neurosci* 11**,** 2881-2895.

Rao, A., and Steward, O. (Year). "Syaptosomal RNA: assessment of contamination by glia and comparison with total RNA", in: *Society for Neuroscience*), 379.

Rao, A., and Steward, O. (1993). Evaluation of RNAs present in synaptodendrosomes: dendritic, glial, and neuronal cell body contribution. *J Neurochem* 61**,** 835-844.

Rapaport, F., Khanin, R., Liang, Y., Pirun, M., Krek, A., Zumbo, P., Mason, C.E., Socci, N.D., and Betel, D. (2013). Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data. *Genome Biol* 14**,** R95.

Rau, A., Gallopin, M., Celeux, G., and Jaffrezic, F. (2013). Data-based filtering for replicated high-throughput transcriptome sequencing experiments. *Bioinformatics* 29**,** 2146-2152.

Ray, D., Kazan, H., Chan, E.T., Pena Castillo, L., Chaudhry, S., Talukder, S., Blencowe, B.J., Morris, Q., and Hughes, T.R. (2009). Rapid and systematic analysis of the RNA recognition specificities of RNA-binding proteins. *Nat Biotechnol* 27**,** 667-670.

Ray, M.T., Shannon Weickert, C., and Webster, M.J. (2014). Decreased BDNF and TrkB mRNA expression in multiple cortical areas of patients with schizophrenia and mood disorders. *Transl Psychiatry* 4**,** e389.

Raymaekers, M., Smets, R., Maes, B., and Cartuyvels, R. (2009). Checklist for optimization and validation of real-time PCR assays. *J Clin Lab Anal* 23**,** 145-151.

Rc, T. (2014). *R: A Language and Environment for Statistical Computing* [Online]. [Accessed].

Redondo, R.L., and Morris, R.G. (2011). Making memories last: the synaptic tagging and capture hypothesis. *Nat Rev Neurosci* 12**,** 17-30.

Reese, A.L., and Kavalali, E.T. (2015). Spontaneous neurotransmission signals through store-driven Ca(2+) transients to maintain synaptic homeostasis. *Elife* 4.

Rhodes, J.S., Best, K., Belknap, J.K., Finn, D.A., and Crabbe, J.C. (2005). Evaluation of a simple model of ethanol drinking to intoxication in C57BL/6J mice. *Physiol Behav* 84**,** 53-63.

Ricciardi, S., Boggio, E.M., Grosso, S., Lonetti, G., Forlani, G., Stefanelli, G., Calcagno, E., Morello, N., Landsberger, N., Biffo, S., Pizzorusso, T., Giustetto, M., and Broccoli, V. (2011). Reduced AKT/mTOR signaling and protein synthesis dysregulation in a Rett syndrome animal model. *Hum Mol Genet* 20**,** 1182-1196.

Rietschel, M., Mattheisen, M., Frank, J., Treutlein, J., Degenhardt, F., Breuer, R., Steffens, M., Mier, D., Esslinger, C., Walter, H., Kirsch, P., Erk, S., Schnell, K., Herms, S., Wichmann, H.E., Schreiber, S., Jockel, K.H., Strohmaier, J., Roeske, D., Haenisch, B., Gross, M., Hoefels, S., Lucae, S., Binder, E.B., Wienker, T.F., Schulze, T.G., Schmal, C., Zimmer, A., Juraeva, D., Brors, B., Bettecken, T., Meyer-Lindenberg, A., Muller-Myhsok, B., Maier, W., Nothen, M.M., and Cichon, S. (2010). Genome-wide association-, replication-, and neuroimaging study implicates HOMER1 in the etiology of major depression. *Biol Psychiatry* 68**,** 578-585.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43**,** e47.

Rizo, J., and Sudhof, T.C. (2002). Snares and Munc18 in synaptic vesicle fusion. *Nat Rev Neurosci* 3**,** 641-653.

Roberts, A., Trapnell, C., Donaghey, J., Rinn, J.L., and Pachter, L. (2011). Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12**,** R22.

Robinson, M.D., Mccarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26**,** 139-140.

Rodriguez, A.J., Shenoy, S.M., Singer, R.H., and Condeelis, J. (2006). Visualization of mRNA translation in living cells. *J Cell Biol* 175**,** 67-76.

Ron, D., and Messing, R.O. (2013). Signaling pathways mediating alcohol effects. *Curr Top Behav Neurosci* 13**,** 87-126.

Rosenberg, T., Gal-Ben-Ari, S., Dieterich, D.C., Kreutz, M.R., Ziv, N.E., Gundelfinger, E.D., and Rosenblum, K. (2014). The roles of protein expression in synaptic plasticity and memory consolidation. *Front Mol Neurosci* 7**,** 86.

Roux, K.J., Kim, D.I., Raida, M., and Burke, B. (2012). A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J Cell Biol* 196**,** 801-810.

Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., Mcgrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., and Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 163**,** 1905-1917.

Sadasivan, S., Zhang, Z., Larner, S.F., Liu, M.C., Zheng, W., Kobeissy, F.H., Hayes, R.L., and Wang, K.K. (2010). Acute NMDA toxicity in cultured rat cerebellar granule neurons is accompanied by autophagy induction and late onset autophagic cell death phenotype. *BMC Neurosci* 11**,** 21.

Sahin, M. (2012). Targeted treatment trials for tuberous sclerosis and autism: no longer a dream. *Curr Opin Neurobiol* 22**,** 895-901.

Sakaida, M., Sukeno, M., Imoto, Y., Tsuchiya, S., Sugimoto, Y., Okuno, Y., and Segi-Nishida, E. (2013). Electroconvulsive seizure-induced changes in gene expression in the mouse hypothalamic paraventricular nucleus. *J Psychopharmacol* 27**,** 1058-1069.

Santini, E., Huynh, T.N., and Klann, E. (2014). Mechanisms of translation control underlying long-lasting synaptic plasticity and the consolidation of long-term memory. *Prog Mol Biol Transl Sci* 122**,** 131-167.

Santini, E., and Klann, E. (2011). Dysregulated mTORC1-Dependent Translational Control: From Brain Disorders to Psychoactive Drugs. *Front Behav Neurosci* 5**,** 76.

Sanz, E., Yang, L., Su, T., Morris, D.R., Mcknight, G.S., and Amieux, P.S. (2009). Cell-type-specific isolation of ribosome-associated mRNA from complex tissues. *Proc Natl Acad Sci U S A* 106**,** 13939-13944.

Sato, A. (2016). mTOR, a Potential Target to Treat Autism Spectrum Disorder. *CNS Neurol Disord Drug Targets* 15**,** 533-543.

Sawicka, K., and Zukin, R.S. (2012). Dysregulation of mTOR signaling in neuropsychiatric disorders: therapeutic implications. *Neuropsychopharmacology* 37**,** 305-306.

Scheckel, C., Drapeau, E., Frias, M.A., Park, C.Y., Fak, J., Zucker-Scharff, I., Kou, Y., Haroutunian, V., Ma'ayan, A., Buxbaum, J.D., and Darnell, R.B. (2016). Regulatory consequences of neuronal ELAV-like protein binding to coding and non-coding RNAs in human brain. *Elife* 5.

Schmidt, E.K., Clavarino, G., Ceppi, M., and Pierre, P. (2009). SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods* 6**,** 275-277.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M., and Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7**,** 3.

Schulte, U., Thumfart, J.O., Klocker, N., Sailer, C.A., Bildl, W., Biniossek, M., Dehn, D., Deller, T., Eble, S., Abbass, K., Wangler, T., Knaus, H.G., and Fakler, B. (2006). The epilepsy-linked Lgi1 protein assembles into presynaptic Kv1 channels and inhibits inactivation by Kvbeta1. *Neuron* 49**,** 697-706.

Searle, B.C. (2010). Scaffold: a bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* 10**,** 1265-1269.

Selak, S., Paternain, A.V., Aller, M.I., Pico, E., Rivera, R., and Lerma, J. (2009). A role for SNAP25 in internalization of kainate receptors and synaptic plasticity. *Neuron* 63**,** 357-371.

Sephton, C.F., Cenik, C., Kucukural, A., Dammer, E.B., Cenik, B., Han, Y., Dewey, C.M., Roth, F.P., Herz, J., Peng, J., Moore, M.J., and Yu, G. (2011). Identification of neuronal RNA targets of TDP-43-containing ribonucleoprotein complexes. *J Biol Chem* 286**,** 1204-1215.

Sephton, C.F., Tang, A.A., Kulkarni, A., West, J., Brooks, M., Stubblefield, J.J., Liu, Y., Zhang, M.Q., Green, C.B., Huber, K.M., Huang, E.J., Herz, J., and Yu, G. (2014). Activity-dependent FUS dysregulation disrupts synaptic homeostasis. *Proc Natl Acad Sci U S A* 111**,** E4769-4778.

Sephton, C.F., and Yu, G. (2015). The function of RNA-binding proteins at the synapse: implications for neurodegeneration. *Cell Mol Life Sci* 72**,** 3621-3635.

Serchov, T., Heumann, R., Van Calker, D., and Biber, K. (2016). Signaling pathways regulating Homer1a expression: implications for antidepressant therapy. *Biol Chem* 397**,** 207-214.

Seto, B. (2012). Rapamycin and mTOR: a serendipitous discovery and implications for breast cancer. *Clin Transl Med* 1**,** 29.

Seyednasrollah, F., Laiho, A., and Elo, L.L. (2015). Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief Bioinform* 16**,** 59-70.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13**,** 2498-2504.

Sharma, A., Hoeffer, C.A., Takayasu, Y., Miyawaki, T., Mcbride, S.M., Klann, E., and Zukin, R.S. (2010). Dysregulation of mTOR signaling in fragile X syndrome. *J Neurosci* 30**,** 694-702.

Sharma, K., Schmitt, S., Bergner, C.G., Tyanova, S., Kannaiyan, N., Manrique-Hoyos, N., Kongi, K., Cantuti, L., Hanisch, U.K., Philips, M.A., Rossner, M.J., Mann, M., and Simons, M. (2015). Cell type- and brain region-resolved mouse brain proteome. *Nat Neurosci* 18**,** 1819-1831.

Shelton, R.C., Claiborne, J., Sidoryk-Wegrzynowicz, M., Reddy, R., Aschner, M., Lewis, D.A., and Mirnics, K. (2011). Altered expression of genes involved in inflammation and apoptosis in frontal cortex in major depression. *Mol Psychiatry* 16**,** 751-762.

Shiraishi-Yamaguchi, Y., and Furuichi, T. (2007). The Homer family proteins. *Genome Biol* 8**,** 206.

Shiraishi, Y., Mizutani, A., Mikoshiba, K., and Furuichi, T. (2003). Coincidence in dendritic clustering and synaptic targeting of homer proteins and NMDA receptor complex proteins NR2B and PSD95 during development of cultured hippocampal neurons. *Mol Cell Neurosci* 22**,** 188-201.

Shu, X., Lev-Ram, V., Deerinck, T.J., Qi, Y., Ramko, E.B., Davidson, M.W., Jin, Y., Ellisman, M.H., and Tsien, R.Y. (2011). A genetically encoded tag for correlated light and electron microscopy of intact cells, tissues, and organisms. *PLoS Biol* 9**,** e1001041.

Shyn, S.I., Shi, J., Kraft, J.B., Potash, J.B., Knowles, J.A., Weissman, M.M., Garriock, H.A., Yokoyama, J.S., Mcgrath, P.J., Peters, E.J., Scheftner, W.A., Coryell, W., Lawson, W.B., Jancic, D., Gejman, P.V., Sanders, A.R., Holmans, P., Slager, S.L., Levinson, D.F., and Hamilton, S.P. (2011). Novel loci for major depression identified by genome-wide association study of Sequenced Treatment Alternatives to Relieve Depression and meta-analysis of three studies. *Mol Psychiatry* 16**,** 202-215.

Si, Y., Liu, P., Li, P., and Brutnell, T.P. (2014). Model-based clustering for RNA-seq data. *Bioinformatics* 30**,** 197-205.

Sigurgeirsson, B., Emanuelsson, O., and Lundeberg, J. (2014). Sequencing degraded RNA addressed by 3' tag counting. *PLoS One* 9**,** e91851.

Silverman, I.M., Li, F., Alexander, A., Goff, L., Trapnell, C., Rinn, J.L., and Gregory, B.D. (2014). RNase-mediated protein footprint sequencing reveals protein-binding sites throughout the human transcriptome. *Genome Biol* 15**,** R3.

Sjostedt, E., Fagerberg, L., Hallstrom, B.M., Haggmark, A., Mitsios, N., Nilsson, P., Ponten, F., Hokfelt, T., Uhlen, M., and Mulder, J. (2015). Defining the Human Brain Proteome Using Transcriptomics and Antibody-Based Profiling with a Focus on the Cerebral Cortex. *PLoS One* 10**,** e0130028.

Skolnick, P., Popik, P., and Trullas, R. (2009). Glutamate-based antidepressants: 20 years on. *Trends Pharmacol Sci* 30**,** 563-569.

Skoog, I., Waern, M., Duberstein, P., Blennow, K., Zetterberg, H., Borjesson-Hanson, A., Ostling, S., Guo, X., Kern, J., Gustafson, D., Gudmundsson, P., Marlow, T., and Kern, S. (2015). A 9-year prospective population-based study on the association

between the APOE*E4 allele and late-life depression in Sweden. *Biol Psychiatry* 78**,** 730-736.

Slonim, D.K., and Yanai, I. (2009). Getting started in gene expression microarray analysis. *PLoS Comput Biol* 5**,** e1000543.

Smalheiser, N.R. (2014). The RNA-centred view of the synapse: non-coding RNAs and synaptic plasticity. *Philos Trans R Soc Lond B Biol Sci* 369.

Smedley, D., Haider, S., Durinck, S., Pandini, L., Provero, P., Allen, J., Arnaiz, O., Awedh, M.H., Baldock, R., Barbiera, G., Bardou, P., Beck, T., Blake, A., Bonierbale, M., Brookes, A.J., Bucci, G., Buetti, I., Burge, S., Cabau, C., Carlson, J.W., Chelala, C., Chrysostomou, C., Cittaro, D., Collin, O., Cordova, R., Cutts, R.J., Dassi, E., Di Genova, A., Djari, A., Esposito, A., Estrella, H., Eyras, E., Fernandez-Banet, J., Forbes, S., Free, R.C., Fujisawa, T., Gadaleta, E., Garcia-Manteiga, J.M., Goodstein, D., Gray, K., Guerra-Assuncao, J.A., Haggarty, B., Han, D.J., Han, B.W., Harris, T., Harshbarger, J., Hastings, R.K., Hayes, R.D., Hoede, C., Hu, S., Hu, Z.L., Hutchins, L., Kan, Z., Kawaji, H., Keliet, A., Kerhornou, A., Kim, S., Kinsella, R., Klopp, C., Kong, L., Lawson, D., Lazarevic, D., Lee, J.H., Letellier, T., Li, C.Y., Lio, P., Liu, C.J., Luo, J., Maass, A., Mariette, J., Maurel, T., Merella, S., Mohamed, A.M., Moreews, F., Nabihoudine, I., Ndegwa, N., Noirot, C., Perez-Llamas, C., Primig, M., Quattrone, A., Quesneville, H., Rambaldi, D., Reecy, J., Riba, M., Rosanoff, S., Saddiq, A.A., Salas, E., Sallou, O., Shepherd, R., Simon, R., Sperling, L., Spooner, W., Staines, D.M., Steinbach, D., Stone, K., Stupka, E., Teague, J.W., Dayem Ullah, A.Z., Wang, J., Ware, D., et al. (2015). The BioMart community portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 43**,** W589-598.

Smith, A.C., Blackshaw, J.A., and Robinson, A.J. (2012). MitoMiner: a data warehouse for mitochondrial proteomics data. *Nucleic Acids Res* 40**,** D1160-1167.

Smith, L.N., Jedynak, J.P., Fontenot, M.R., Hale, C.F., Dietz, K.C., Taniguchi, M., Thomas, F.S., Zirlin, B.C., Birnbaum, S.G., Huber, K.M., Thomas, M.J., and Cowan, C.W. (2014). Fragile X mental retardation protein regulates synaptic and behavioral plasticity to repeated cocaine administration. *Neuron* 82**,** 645-658.

Sneddon, M.W., Faeder, J.R., and Emonet, T. (2011). Efficient modeling, simulation and coarse-graining of biological complexity with NFsim. *Nat Methods* 8**,** 177-183.

Sobczyk, A., Scheuss, V., and Svoboda, K. (2005). NMDA receptor subunit-dependent [Ca2+] signaling in individual hippocampal dendritic spines. *J Neurosci* 25**,** 6037-6046.

Soden, M.E., and Chen, L. (2010). Fragile X protein FMRP is required for homeostatic plasticity and regulation of synaptic strength by retinoic acid. *J Neurosci* 30**,** 16910-16921.

Soneson, C., and Delorenzi, M. (2013). A comparison of methods for differential expression analysis of RNA-seq data. *BMC Bioinformatics* 14**,** 91.

Sorokina, O., Sorokin, A., and Armstrong, J.D. (2011). Towards a quantitative model of the post-synaptic proteome. *Mol Biosyst* 7**,** 2813-2823.

Sosanya, N.M., Brager, D.H., Wolfe, S., Niere, F., and Raab-Graham, K.F. (2015a). Rapamycin reveals an mTOR-independent repression of Kv1.1 expression during epileptogenesis. *Neurobiol Dis* 73**,** 96-105.

Sosanya, N.M., Cacheaux, L.P., Workman, E.R., Niere, F., Perrone-Bizzozero, N.I., and Raab-Graham, K.F. (2015b). Mammalian Target of Rapamycin (mTOR) Tagging Promotes Dendritic Branch Variability through the Capture of Ca2+/Calmodulin-dependent Protein Kinase II alpha (CaMKIIalpha) mRNAs by the RNA-binding Protein HuD. *J Biol Chem* 290**,** 16357-16371.

Sosanya, N.M., Huang, P.P., Cacheaux, L.P., Chen, C.J., Nguyen, K., Perrone-Bizzozero, N.I., and Raab-Graham, K.F. (2013). Degradation of high affinity HuD targets releases Kv1.1 mRNA from miR-129 repression by mTORC1. *J Cell Biol* 202**,** 53-69.

Spilman, P., Podlutskaya, N., Hart, M.J., Debnath, J., Gorostiza, O., Bredesen, D., Richardson, A., Strong, R., and Galvan, V. (2010). Inhibition of mTOR by rapamycin abolishes cognitive deficits and reduces amyloid-beta levels in a mouse model of Alzheimer's disease. *PLoS One* 5**,** e9979.

Spitzer, J., Hafner, M., Landthaler, M., Ascano, M., Farazi, T., Wardle, G., Nusbaum, J., Khorshid, M., Burger, L., Zavolan, M., and Tuschl, T. (2014). PAR-CLIP (Photoactivatable Ribonucleoside-Enhanced Crosslinking and Immunoprecipitation): a step-by-step protocol to the transcriptome-wide identification of binding sites of RNA-binding proteins. *Methods Enzymol* 539**,** 113-161.

Stachowicz, A., Glombik, K., Olszanecki, R., Basta-Kaim, A., Suski, M., Lason, W., and Korbut, R. (2016). The impact of mitochondrial aldehyde dehydrogenase (ALDH2) activation by Alda-1 on the behavioral and biochemical disturbances in animal model of depression. *Brain Behav Immun* 51**,** 144-153.

Stambolic, V., Macpherson, D., Sas, D., Lin, Y., Snow, B., Jang, Y., Benchimol, S., and Mak, T.W. (2001). Regulation of PTEN transcription by p53. *Mol Cell* 8**,** 317-325.

Steward, O., and Falk, P.M. (1985). Polyribosomes under developing spine synapses: growth specializations of dendrites at sites of synaptogenesis. *J Neurosci Res* 13**,** 75-88.

Steward, O., and Falk, P.M. (1986). Protein-synthetic machinery at postsynaptic sites during synaptogenesis: a quantitative study of the association between polyribosomes and developing synapses. *J Neurosci* 6**,** 412-423.

Steward, O., Falk, P.M., and Torre, E.R. (1996). Ultrastructural basis for gene expression at the synapse: synapse-associated polyribosome complexes. *J Neurocytol* 25**,** 717-734.

Steward, O., and Levy, W.B. (1982). Preferential localization of polyribosomes under the base of dendritic spines in granule cells of the dentate gyrus. *J Neurosci* 2**,** 284-291.

Steward, O., and Schuman, E.M. (2001). Protein synthesis at synaptic sites on dendrites. *Annu Rev Neurosci* 24**,** 299-325.

Stewart, M.G., and Popov, V.I. (eds.). (2012). *Stuctural synaptic and dendritic spine plasticity in the hippocampus.* Signapore: World Scientific Publishing.

Stoica, L., Zhu, P.J., Huang, W., Zhou, H., Kozma, S.C., and Costa-Mattioli, M. (2011). Selective pharmacogenetic inhibition of mammalian target of Rapamycin complex I (mTORC1) blocks long-term synaptic plasticity and memory storage. *Proc Natl Acad Sci U S A* 108**,** 3791-3796.

Su, S.C., and Tsai, L.H. (2011). Cyclin-dependent kinases in brain development and disease. *Annu Rev Cell Dev Biol* 27**,** 465-491.

Sugimoto, Y., Konig, J., Hussain, S., Zupan, B., Curk, T., Frye, M., and Ule, J. (2012). Analysis of CLIP and iCLIP methods for nucleotide-resolution studies of protein-RNA interactions. *Genome Biol* 13**,** R67.

Sugimoto, Y., Vigilante, A., Darbo, E., Zirra, A., Militti, C., D'ambrogio, A., Luscombe, N.M., and Ule, J. (2015). hiCLIP reveals the in vivo atlas of mRNA secondary structures recognized by Staufen 1. *Nature* 519**,** 491-494.

Suhl, J.A., Chopra, P., Anderson, B.R., Bassell, G.J., and Warren, S.T. (2014). Analysis of FMRP mRNA target datasets reveals highly associated mRNAs mediated by G-quadruplex structures formed via clustered WGGA sequences. *Hum Mol Genet* 23**,** 5479-5491.

Sun, Y., Fang, Y., Yoon, M.S., Zhang, C., Roccio, M., Zwartkruis, F.J., Armstrong, M., Brown, H.A., and Chen, J. (2008). Phospholipase D1 is an effector of Rheb in the mTOR pathway. *Proc Natl Acad Sci U S A* 105**,** 8286-8291.

Surget, A., Saxe, M., Leman, S., Ibarguen-Vargas, Y., Chalon, S., Griebel, G., Hen, R., and Belzung, C. (2008). Drug-dependent requirement of hippocampal neurogenesis in a model of depression and of antidepressant reversal. *Biol Psychiatry* 64**,** 293-301.

Sutherland, B.W., Toews, J., and Kast, J. (2008). Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. *J Mass Spectrom* 43**,** 699-715.

Sutton, M.A., Ito, H.T., Cressy, P., Kempf, C., Woo, J.C., and Schuman, E.M. (2006). Miniature neurotransmission stabilizes synaptic function via tonic suppression of local dendritic protein synthesis. *Cell* 125**,** 785-799.

Sutton, M.A., and Schuman, E.M. (2006). Dendritic protein synthesis, synaptic plasticity, and memory. *Cell* 127**,** 49-58.

Sutton, M.A., Wall, N.R., Aakalu, G.N., and Schuman, E.M. (2004). Regulation of dendritic protein synthesis by miniature synaptic events. *Science* 304**,** 1979-1983.

Swanger, S.A., Bassell, G.J., and Gross, C. (2011). High-resolution fluorescence in situ hybridization to detect mRNAs in neuronal compartments in vitro and in vivo. *Methods Mol Biol* 714**,** 103-123.

Swiech, L., Perycz, M., Malik, A., and Jaworski, J. (2008). Role of mTOR in physiology and pathology of the nervous system. *Biochim Biophys Acta* 1784**,** 116-132.

Szumlinski, K.K., Lominac, K.D., Kleschen, M.J., Oleson, E.B., Dehoff, M.H., Schwarz, M.K., Seeburg, P.H., Worley, P.F., and Kalivas, P.W. (2005). Behavioral and neurochemical phenotyping of Homer1 mutant mice: possible relevance to schizophrenia. *Genes Brain Behav* 4**,** 273-288.

Tadic, A., Rujescu, D., Muller, M.J., Kohnen, R., Stassen, H.H., Dahmen, N., and Szegedi, A. (2007). A monoamine oxidase B gene variant and short-term antidepressant treatment response. *Prog Neuropsychopharmacol Biol Psychiatry* 31**,** 1370-1377.

Takahara, T., and Maeda, T. (2012). Stress granules: the last refuge of TORC1? *Cell Cycle* 11**,** 3707-3708.

Takei, N., Inamura, N., Kawamura, M., Namba, H., Hara, K., Yonezawa, K., and Nawa, H. (2004). Brain-derived neurotrophic factor induces mammalian target of rapamycin-dependent local activation of translation machinery and protein synthesis in neuronal dendrites. *J Neurosci* 24**,** 9760-9769.

Tan, K.M., Peterson, A., and Witten, D. (eds.). (2014). *Classification of RNAseq data.* Switzerland: Springer.

Tanenbaum, M.E., Gilbert, L.A., Qi, L.S., Weissman, J.S., and Vale, R.D. (2014). A protein-tagging system for signal amplification in gene expression and fluorescence imaging. *Cell* 159**,** 635-646.

Tang, S.J., Meulemans, D., Vazquez, L., Colaco, N., and Schuman, E. (2001). A role for a rat homolog of staufen in the transport of RNA to neuronal dendrites. *Neuron* 32**,** 463-475.

Tang, S.J., Reis, G., Kang, H., Gingras, A.C., Sonenberg, N., and Schuman, E.M. (2002). A rapamycin-sensitive signaling pathway contributes to long-term synaptic plasticity in the hippocampus. *Proc Natl Acad Sci U S A* 99**,** 467-472.

Tatavarty, V., Ifrim, M.F., Levin, M., Korza, G., Barbarese, E., Yu, J., and Carson, J.H. (2012). Single-molecule imaging of translational output from individual RNA granules in neurons. *Mol Biol Cell* 23**,** 918-929.

Tavazoie, S.F., Alvarez, V.A., Ridenour, D.A., Kwiatkowski, D.J., and Sabatini, B.L. (2005). Regulation of neuronal morphology and function by the tumor suppressors Tsc1 and Tsc2. *Nat Neurosci* 8**,** 1727-1734.

Terunuma, M., Vargas, K.J., Wilkins, M.E., Ramirez, O.A., Jaureguiberry-Bravo, M., Pangalos, M.N., Smart, T.G., Moss, S.J., and Couve, A. (2010). Prolonged activation of NMDA receptors promotes dephosphorylation and alters postendocytic sorting of GABAB receptors. *Proc Natl Acad Sci U S A* 107**,** 13918-13923.

Thomas, M.G., Pascual, M.L., Maschi, D., Luchelli, L., and Boccaccio, G.L. (2014). Synaptic control of local translation: the plot thickens with new characters. *Cell Mol Life Sci* 71**,** 2219-2239.

Thorvaldsdottir, H., Robinson, J.T., and Mesirov, J.P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 14**,** 178-192.

Tom Dieck, S., Kochen, L., Hanus, C., Heumuller, M., Bartnik, I., Nassim-Assir, B., Merk, K., Mosler, T., Garg, S., Bunse, S., Tirrell, D.A., and Schuman, E.M. (2015). Direct visualization of newly synthesized target proteins in situ. *Nat Methods* 12**,** 411-414.

Tom Dieck, S., Muller, A., Nehring, A., Hinz, F.I., Bartnik, I., Schuman, E.M., and Dieterich, D.C. (2012). Metabolic labeling with noncanonical amino acids and

visualization by chemoselective fluorescent tagging. *Curr Protoc Cell Biol* Chapter 7**,** Unit7 11.

Tomasoni, R., Repetto, D., Morini, R., Elia, C., Gardoni, F., Di Luca, M., Turco, E., Defilippi, P., and Matteoli, M. (2013). SNAP-25 regulates spine formation through postsynaptic binding to p140Cap. *Nat Commun* 4**,** 2136.

Torre, E.R., and Steward, O. (1992). Demonstration of local protein synthesis within dendrites using a new cell culture system that permits the isolation of living axons and dendrites from their cell bodies. *J Neurosci* 12**,** 762-772.

Trapnell, C., Hendrickson, D.G., Sauvageau, M., Goff, L., Rinn, J.L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat Biotechnol* 31**,** 46-53.

Treit, D., and Fundytus, M. (1988). Thigmotaxis as a test for anxiolytic activity in rats. *Pharmacol Biochem Behav* 31**,** 959-962.

Trivedi, M.H., Hollander, E., Nutt, D., and Blier, P. (2008). Clinical evidence and potential neurobiological underpinnings of unresolved symptoms of depression. *J Clin Psychiatry* 69**,** 246-258.

Trivedi, M.H., Rush, A.J., Wisniewski, S.R., Nierenberg, A.A., Warden, D., Ritz, L., Norquist, G., Howland, R.H., Lebowitz, B., Mcgrath, P.J., Shores-Wilson, K., Biggs, M.M., Balasubramani, G.K., Fava, M., and Team, S.D.S. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 163**,** 28-40.

Tsokas, P., Grace, E.A., Chan, P., Ma, T., Sealfon, S.C., Iyengar, R., Landau, E.M., and Blitzer, R.D. (2005). Local protein synthesis mediates a rapid increase in dendritic elongation factor 1A after induction of late long-term potentiation. *J Neurosci* 25**,** 5833-5843.

Tu, J.C., Xiao, B., Naisbitt, S., Yuan, J.P., Petralia, R.S., Brakeman, P., Doan, A., Aakalu, V.K., Lanahan, A.A., Sheng, M., and Worley, P.F. (1999). Coupling of mGluR/Homer and PSD-95 complexes by the Shank family of postsynaptic density proteins. *Neuron* 23**,** 583-592.

Turner, E.H., Matthews, A.M., Linardatos, E., Tell, R.A., and Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 358**,** 252-260.

Turrigiano, G. (2012). Homeostatic synaptic plasticity: local and global mechanisms for stabilizing neuronal function. *Cold Spring Harb Perspect Biol* 4**,** a005736.

Turrigiano, G.G. (2008). The self-tuning neuron: synaptic scaling of excitatory synapses. *Cell* 135**,** 422-435.

Turrigiano, G.G., and Nelson, S.B. (2000). Hebb and homeostasis in neuronal plasticity. *Curr Opin Neurobiol* 10**,** 358-364.

Ule, J., Jensen, K., Mele, A., and Darnell, R.B. (2005). CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods* 37**,** 376-386.

Ule, J., Jensen, K.B., Ruggiu, M., Mele, A., Ule, A., and Darnell, R.B. (2003). CLIP identifies Nova-regulated RNA networks in the brain. *Science* 302**,** 1212-1215.

Undurraga, J., and Baldessarini, R.J. (2012). Randomized, placebo-controlled trials of antidepressants for acute major depression: thirty-year meta-analytic review. *Neuropsychopharmacology* 37**,** 851-864.

Unschuld, P.G., Ising, M., Specht, M., Erhardt, A., Ripke, S., Heck, A., Kloiber, S., Straub, V., Brueckl, T., Muller-Myhsok, B., Holsboer, F., and Binder, E.B. (2009). Polymorphisms in the GAD2 gene-region are associated with susceptibility for unipolar depression and with a risk factor for anxiety disorders. *Am J Med Genet B Neuropsychiatr Genet* 150B**,** 1100-1109.

Valor, L.M., and Grant, S.G. (2007). Integrating synapse proteomics with transcriptional regulation. *Behav Genet* 37**,** 18-30.

Van Der Brug, M.P., Blackinton, J., Chandran, J., Hao, L.Y., Lal, A., Mazan-Mamczarz, K., Martindale, J., Xie, C., Ahmad, R., Thomas, K.J., Beilina, A., Gibbs, J.R., Ding, J., Myers, A.J., Zhan, M., Cai, H., Bonini, N.M., Gorospe, M., and Cookson, M.R. (2008). RNA binding activity of the recessive parkinsonism protein DJ-1 supports involvement in multiple cellular pathways. *Proc Natl Acad Sci U S A* 105**,** 10244-10249.

Van Dijk, E.L., Jaszczyszyn, Y., and Thermes, C. (2014). Library preparation methods for next-generation sequencing: tone down the bias. *Exp Cell Res* 322**,** 12-20.

Vander Haar, E., Lee, S.I., Bandhakavi, S., Griffin, T.J., and Kim, D.H. (2007). Insulin signalling to mTOR mediated by the Akt/PKB substrate PRAS40. *Nat Cell Biol* 9**,** 316-323.

Verghese, P.B., Castellano, J.M., and Holtzman, D.M. (2011). Apolipoprotein E in Alzheimer's disease and other neurological disorders. *Lancet Neurol* 10**,** 241-252.

Verity, M.A., Brown, W.J., and Cheung, M. (1980). Isolation of ribosome containing synaptosome subpopulation with active in vitro protein synthesis. *J Neurosci Res* 5**,** 143-153.

Vermeulen, J., De Preter, K., Lefever, S., Nuytens, J., De Vloed, F., Derveaux, S., Hellemans, J., Speleman, F., and Vandesompele, J. (2011). Measurable impact of RNA quality on gene expression results from quantitative PCR. *Nucleic Acids Res* 39**,** e63.

Vijay, N., Poelstra, J.W., Kunstner, A., and Wolf, J.B. (2013). Challenges and strategies in transcriptome assembly and differential gene expression quantification. A comprehensive in silico assessment of RNA-seq experiments. *Mol Ecol* 22**,** 620-634.

Villasana, L.E., Klann, E., and Tejada-Simon, M.V. (2006). Rapid isolation of synaptoneurosomes and postsynaptic densities from adult mouse hippocampus. *J Neurosci Methods* 158**,** 30-36.

Vincent, P., and Mulle, C. (2009). Kainate receptors in epilepsy and excitotoxicity. *Neuroscience* 158**,** 309-323.

Vizcaino, J.A., Csordas, A., Del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., Xu, Q.W., Wang, R., and Hermjakob, H. (2016). 2016 update of the PRIDE database and its related tools. *Nucleic Acids Res* 44**,** 11033.

Vizcaino, J.A., Deutsch, E.W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J.A., Sun, Z., Farrah, T., Bandeira, N., Binz, P.A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R.J., Kraus, H.J., Albar, J.P., Martinez-Bartolome, S., Apweiler, R., Omenn, G.S., Martens, L., Jones, A.R., and Hermjakob, H. (2014). ProteomeXchange provides globally coordinated proteomics data submission and dissemination. *Nat Biotechnol* 32**,** 223-226.

Wagner, K.V., Hartmann, J., Labermaier, C., Hausl, A.S., Zhao, G., Harbich, D., Schmid, B., Wang, X.D., Santarelli, S., Kohl, C., Gassen, N.C., Matosin, N., Schieven, M., Webhofer, C., Turck, C.W., Lindemann, L., Jaschke, G., Wettstein, J.G., Rein, T., Muller, M.B., and Schmidt, M.V. (2015). Homer1/mGluR5 activity moderates vulnerability to chronic social stress. *Neuropsychopharmacology* 40**,** 1222-1233.

Wagnon, J.L., Briese, M., Sun, W., Mahaffey, C.L., Curk, T., Rot, G., Ule, J., and Frankel, W.N. (2012). CELF4 regulates translation and local abundance of a vast set of mRNAs, including genes associated with regulation of synaptic function. *PLoS Genet* 8**,** e1003067.

Wallace, W., and Bear, M.F. (2004). A morphological correlate of synaptic scaling in visual cortex. *J Neurosci* 24**,** 6928-6938.

Wang, D.O., Kim, S.M., Zhao, Y., Hwang, H., Miura, S.K., Sossin, W.S., and Martin, K.C. (2009). Synapse- and stimulus-specific local translation during long-term neuronal plasticity. *Science* 324**,** 1536-1540.

Wang, H., and Doering, L.C. (2013). Reversing autism by targeting downstream mTOR signaling. *Front Cell Neurosci* 7**,** 28.

Wang, H., and Song, M. (2011). Ckmeans.1d.dp: Optimal k-means clustering in one dimension by dynamic programming. *The R Journal* 320**,** 5.

Wang, L., Feng, Z., Wang, X., Wang, X., and Zhang, X. (2010a). DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* 26**,** 136-138.

Wang, L., Wang, S., and Li, W. (2012a). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28**,** 2184-2185.

Wang, S., Wu, J., Nie, S.D., Bereczki, E., and Pei, J.J. (2013). Dysregulated mTOR-dependent signaling in neurodegeneration or carcinogenesis: implication for Alzheimer's disease and brain tumors. *J Alzheimers Dis* 37**,** 495-505.

Wang, T., Xiao, G., Chu, Y., Zhang, M.Q., Corey, D.R., and Xie, Y. (2015). Design and bioinformatics analysis of genome-wide CLIP experiments. *Nucleic Acids Res* 43**,** 5263-5274.

Wang, W., Kwon, E.J., and Tsai, L.H. (2012b). MicroRNAs in learning, memory, and neurological diseases. *Learn Mem* 19**,** 359-368.

Wang, X., and Proud, C.G. (2011). mTORC1 signaling: what we still don't know. *J Mol Cell Biol* 3**,** 206-220.

Wang, Y., Ghaffari, N., Johnson, C.D., Braga-Neto, U.M., Wang, H., Chen, R., and Zhou, H. (2011). Evaluation of the coverage and depth of transcriptome by RNA-Seq in chickens. *BMC Bioinformatics* 12 Suppl 10**,** S5.

Wang, Z., Kayikci, M., Briese, M., Zarnack, K., Luscombe, N.M., Rot, G., Zupan, B., Curk, T., and Ule, J. (2010b). iCLIP predicts the dual splicing effects of TIA-RNA interactions. *PLoS Biol* 8**,** e1000530.

Warner-Schmidt, J.L., and Duman, R.S. (2006). Hippocampal neurogenesis: opposing effects of stress and antidepressant treatment. *Hippocampus* 16**,** 239-249.

Warnes, G.R., Bolker, B., Gorjanc, G., Grothendieck, G., Korosec, A., Lumley, T., Macqueen, D., Magnusson, A., and Rogers, J. (2015). "gdata: Various R Programming Tools for Data Manipulation.". R package version 2.17.0 ed.).

Washbourne, P., Thompson, P.M., Carta, M., Costa, E.T., Mathews, J.R., Lopez-Bendito, G., Molnar, Z., Becher, M.W., Valenzuela, C.F., Partridge, L.D., and Wilson, M.C. (2002). Genetic ablation of the t-SNARE SNAP-25 distinguishes mechanisms of neuroexocytosis. *Nat Neurosci* 5**,** 19-26.

Waung, M.W., Pfeiffer, B.E., Nosyreva, E.D., Ronesi, J.A., and Huber, K.M. (2008). Rapid translation of Arc/Arg3.1 selectively mediates mGluR-dependent LTD through persistent increases in AMPAR endocytosis rate. *Neuron* 59**,** 84-97.

Weiler, I.J., Irwin, S.A., Klintsova, A.Y., Spencer, C.M., Brazelton, A.D., Miyashiro, K., Comery, T.A., Patel, B., Eberwine, J., and Greenough, W.T. (1997). Fragile X mental retardation protein is translated near synapses in response to neurotransmitter activation. *Proc Natl Acad Sci U S A* 94**,** 5395-5400.

Weston, M.C., Chen, H., and Swann, J.W. (2014). Loss of mTOR repressors Tsc1 or Pten has divergent effects on excitatory and inhibitory synaptic transmission in single hippocampal neuron cultures. *Front Mol Neurosci* 7**,** 1.

Weyn-Vanhentenryck, S.M., Mele, A., Yan, Q., Sun, S., Farny, N., Zhang, Z., Xue, C., Herre, M., Silver, P.A., Zhang, M.Q., Krainer, A.R., Darnell, R.B., and Zhang, C. (2014). HITS-CLIP and integrative modeling define the Rbfox splicing-regulatory network linked to brain development and autism. *Cell Rep* 6**,** 1139-1152.

Whittaker, V.P., Michaelson, I.A., and Kirkland, R.J. (1964). The separation of synaptic vesicles from nerve-ending particles ('synaptosomes'). *Biochem J* 90**,** 293-303.

Wickham, H. (2007). Reshaping Data with the reshape Package. *J Stat Softw* 21**,** 20.

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis.* New York: Springer.

Wickham, H. (2011). The Split-Apply-Combine Strategy for Data Analysis. *J Stat Softw* 40**,** 29.

Wickham, H., and Francois, R. (2016). "dplyr: A Grammar of Data Manipulation". R package version 0.5.0 ed.).

Williams, A.G., Thomas, S., Wyman, S.K., and Holloway, A.K. (2014). RNA-seq Data: Challenges in and Recommendations for Experimental Design and Analysis. *Curr Protoc Hum Genet* 83**,** 11 13 11-20.

Williams, C.R., Baccarella, A., Parrish, J.Z., and Kim, C.C. (2016). Trimming of sequence reads alters RNA-Seq gene expression estimates. *BMC Bioinformatics* 17**,** 103.

Witten, D.M. (2011). Classification and clustering of sequencing data using a poisson model. *Ann Appl Stat* 5**,** 2493-2518.

Wolfe, S.A., Workman, E.R., Heaney, C.F., Niere, F., Namjoshi, S., Cacheaux, L.P., Farris, S.P., Drew, M.R., Zemelman, B.V., Harris, R.A., and Raab-Graham, K.F. (2016).

FMRP regulates an ethanol-dependent shift in GABABR function and expression with rapid antidepressant properties. *Nat Commun* 7**,** 12867.

Wong, M. (2014). mTOR Strikes Again: mTORC1 Activation Causes Epilepsy Independent of Overt Pathological Changes. *Epilepsy Curr* 14**,** 41-43.

Workman, E.R., Haddick, P.C., Bush, K., Dilly, G.A., Niere, F., Zemelman, B.V., and Raab-Graham, K.F. (2015). Rapid antidepressants stimulate the decoupling of GABA(B) receptors from GIRK/Kir3 channels through increased protein stability of 14-3-3eta. *Mol Psychiatry* 20**,** 298-310.

Workman, E.R., Niere, F., and Raab-Graham, K.F. (2013). mTORC1-dependent protein synthesis underlying rapid antidepressant effect requires GABABR signaling. *Neuropharmacology* 73**,** 192-203.

Wu, B., Eliscovich, C., Yoon, Y.J., and Singer, R.H. (2016). Translation dynamics of single mRNAs in live cells and neurons. *Science* 352**,** 1430-1435.

Wykes, R.C., Heeroma, J.H., Mantoan, L., Zheng, K., Macdonald, D.C., Deisseroth, K., Hashemi, K.S., Walker, M.C., Schorge, S., and Kullmann, D.M. (2012). Optogenetic and potassium channel gene therapy in a rodent model of focal neocortical epilepsy. *Sci Transl Med* 4**,** 161ra152.

Xia, Z., and Storm, D.R. (2005). The role of calmodulin as a signal integrator for synaptic plasticity. *Nat Rev Neurosci* 6**,** 267-276.

Xiao, B., Tu, J.C., Petralia, R.S., Yuan, J.P., Doan, A., Breder, C.D., Ruggiero, A., Lanahan, A.A., Wenthold, R.J., and Worley, P.F. (1998). Homer regulates the association of group 1 metabotropic glutamate receptors with multivalent complexes of homer-related, synaptic proteins. *Neuron* 21**,** 707-716.

Xiao, B., Tu, J.C., and Worley, P.F. (2000). Homer: a link between neural activity and glutamate receptor function. *Curr Opin Neurobiol* 10**,** 370-374.

Xie, Y. (2016). "DT: A Wrapper of the JavaScript Library 'DataTables'". R package version 0.2 ed.).

Xin, X.Y., Pan, J., Wang, X.Q., Ma, J.F., Ding, J.Q., Yang, G.Y., and Chen, S.D. (2011). 2-methoxyestradiol attenuates autophagy activation after global ischemia. *Can J Neurol Sci* 38**,** 631-638.

Yashiro, K., and Philpot, B.D. (2008). Regulation of NMDA receptor subunit expression and its implications for LTD, LTP, and metaplasticity. *Neuropharmacology* 55**,** 1081-1094.

Yates, S.C., Zafar, A., Hubbard, P., Nagy, S., Durant, S., Bicknell, R., Wilcock, G., Christie, S., Esiri, M.M., Smith, A.D., and Nagy, Z. (2013). Dysfunction of the mTOR pathway is a risk factor for Alzheimer's disease. *Acta Neuropathol Commun* 1**,** 3.

Yin, J., and Yuan, Q. (2014). Structural homeostasis in the nervous system: a balancing act for wiring plasticity and stability. *Front Cell Neurosci* 8**,** 439.

Yokoi, N., Fukata, M., and Fukata, Y. (2012). Synaptic plasticity regulated by protein-protein interactions and posttranslational modifications. *Int Rev Cell Mol Biol* 297**,** 1-43.

Yoshimasu, K., Mure, K., Hashimoto, M., Takemura, S., Tsuno, K., Hayashida, M., Kinoshita, K., Takeshita, T., and Miyashita, K. (2015). Genetic alcohol sensitivity regulated by ALDH2 and ADH1B polymorphisms is strongly associated with depression and anxiety in Japanese employees. *Drug Alcohol Depend* 147**,** 130-136.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol* 11**,** R14.

Yu, H., Greenbaum, D., Xin Lu, H., Zhu, X., and Gerstein, M. (2004). Genomic analysis of essentiality within protein networks. *Trends Genet* 20**,** 227-231.

Yu, W., Zhu, H., Wang, Y., Li, G., Wang, L., and Li, H. (2015). Reactive Transformation and Increased BDNF Signaling by Hippocampal Astrocytes in Response to MK-801. *PLoS One* 10**,** e0145651.

Zang, J.B., Nosyreva, E.D., Spencer, C.M., Volk, L.J., Musunuru, K., Zhong, R., Stone, E.F., Yuva-Paylor, L.A., Huber, K.M., Paylor, R., Darnell, J.C., and Darnell, R.B. (2009). A mouse model of the human Fragile X syndrome I304N mutation. *PLoS Genet* 5**,** e1000758.

Zarate, C., Duman, R.S., Liu, G., Sartori, S., Quiroz, J., and Murck, H. (2013). New paradigms for treatment-resistant depression. *Ann N Y Acad Sci* 1292**,** 21-31.

Zarate, C., Jr., Machado-Vieira, R., Henter, I., Ibrahim, L., Diazgranados, N., and Salvadore, G. (2010). Glutamatergic modulators: the future of treating mood disorders? *Harv Rev Psychiatry* 18**,** 293-303.

Zarate, C.A., Jr., Singh, J.B., Carlson, P.J., Brutsche, N.E., Ameli, R., Luckenbaugh, D.A., Charney, D.S., and Manji, H.K. (2006). A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry* 63**,** 856-864.

Zeng, L.H., Rensing, N.R., and Wong, M. (2009). The mammalian target of rapamycin signaling pathway mediates epileptogenesis in a model of temporal lobe epilepsy. *J Neurosci* 29**,** 6964-6972.

Zeng, L.H., Rensing, N.R., Zhang, B., Gutmann, D.H., Gambello, M.J., and Wong, M. (2011). Tsc2 gene inactivation causes a more severe epilepsy phenotype than Tsc1 inactivation in a mouse model of tuberous sclerosis complex. *Hum Mol Genet* 20**,** 445-454.

Zhang, C., and Darnell, R.B. (2011). Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat Biotechnol* 29**,** 607-614.

Zhang, W., and Linden, D.J. (2003). The other side of the engram: experience-driven changes in neuronal intrinsic excitability. *Nat Rev Neurosci* 4**,** 885-900.

Zhang, Y., Chen, K., Sloan, S.A., Bennett, M.L., Scholze, A.R., O'keeffe, S., Phatnani, H.P., Guarnieri, P., Caneda, C., Ruderisch, N., Deng, S., Liddelow, S.A., Zhang, C., Daneman, R., Maniatis, T., Barres, B.A., and Wu, J.Q. (2014). An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* 34**,** 11929-11947.

Zhang, Y., Xie, S., Xu, H., and Qu, L. (2015). CLIP: viewing the RNA world from an RNA-protein interactome perspective. *Sci China Life Sci* 58**,** 75-88.

Zhao, W., He, X., Hoadley, K.A., Parker, J.S., Hayes, D.N., and Perou, C.M. (2014). Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. *BMC Genomics* 15**,** 419.

Zhou, H., Mangelsdorf, M., Liu, J., Zhu, L., and Wu, J.Y. (2014). RNA-binding proteins in neurological diseases. *Sci China Life Sci* 57**,** 432-444.

Zhou, J., Blundell, J., Ogawa, S., Kwon, C.H., Zhang, W., Sinton, C., Powell, C.M., and Parada, L.F. (2009a). Pharmacological inhibition of mTORC1 suppresses anatomical, cellular, and behavioral abnormalities in neural-specific Pten knock-out mice. *J Neurosci* 29**,** 1773-1783.

Zhou, Y.D., Lee, S., Jin, Z., Wright, M., Smith, S.E., and Anderson, M.P. (2009b). Arrested maturation of excitatory synapses in autosomal dominant lateral temporal lobe epilepsy. *Nat Med* 15**,** 1208-1214.

Zhou, Z., Yuan, Q., Mash, D.C., and Goldman, D. (2011). Substance-specific and shared transcription and epigenetic changes in the human hippocampus chronically exposed to cocaine and alcohol. *Proc Natl Acad Sci U S A* 108**,** 6626-6631.

Zoghbi, H.Y., and Bear, M.F. (2012). Synaptic dysfunction in neurodevelopmental disorders associated with autism and intellectual disabilities. *Cold Spring Harb Perspect Biol* 4.

Zukin, R.S., Richter, J.D., and Bagni, C. (2009). Signals, synapses, and synthesis: how new proteins control plasticity. *Front Neural Circuits* 3**,** 14.

Zuo, Y., Yang, G., Kwon, E., and Gan, W.B. (2005). Long-term sensory deprivation prevents dendritic spine loss in primary somatosensory cortex. *Nature* 436**,** 261-265.

Zwiener, I., Frisch, B., and Binder, H. (2014). Transforming RNA-Seq data to improve the performance of prognostic gene signatures. *PLoS One* 9**,** e85150.

Zyprych-Walczak, J., Szabelska, A., Handschuh, L., Gorczak, K., Klamecka, K., Figlerowicz, M., and Siatkowski, I. (2015). The Impact of Normalization Methods on RNA-Seq Data Analysis. *Biomed Res Int* 2015**,** 621690.